

HIGH-SPEED WHITE LIGHT INTERFEROMETRY FOR IMAGING APPLICATIONS

by

Emma Catherine Landsiedel

---

Copyright © Emma Catherine Landsiedel 2019

A Thesis Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

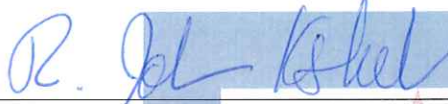
In the Graduate College

THE UNIVERSITY OF ARIZONA

2019

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis prepared by Emma Landsiedel, titled *High-Speed White Light Interferometry for Imaging Applications* and recommend that it be accepted as fulfilling the dissertation requirement for the Master's Degree.

  
\_\_\_\_\_  
R. John Koshel

Date: 4/22/19

  
\_\_\_\_\_  
Amit Ashok

Date: 4/22/2019

  
\_\_\_\_\_  
Dae Wook Kim

Date: 4/22/2019

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

I hereby certify that I have read this thesis prepared under my direction and recommend that it be accepted as fulfilling the Master's requirement.

  
\_\_\_\_\_  
R. John Koshel  
Master's Thesis Committee Chair  
College of Optical Sciences

Date: 4/22/19

ARIZONA

## Acknowledgements

To Dr. John Koshel: Thank you for your guidance through this project as my advisor. I am appreciative of receiving funding to do this interesting and challenging work, and of having an encouraging, engaged, and committed advisor to support me through it all. Your questions and feedback have helped me grow as an optical engineer, and I have greatly enjoyed collaborating with you on this research.

To the College of Optical Sciences: Thank you for becoming the place I call home. Thank you for welcoming me into your tight-knit community and supporting me through my BS and MS at the University of Arizona. I am glad to have had professors to challenge and inspire me, and a friendly and funny academic programs staff to help me. I am also so thankful for my fellow classmates who have become some of my closest and most trusted friends.

To my family: Thank you for everything. There aren't enough words to say all you have done for me, so it comes down to this: thank you for math problems at the breakfast table. Thank you for building popsicle stick bridges with me. Thank you for your support, love, and interest in what I do. Thank you for being my first and longest-lasting teachers – I would not be where I am without you.

# Table of Contents

List of Figures .....	6
List of Tables .....	9
Abstract .....	10
1. Introduction .....	11
1.1 Types of EDOF systems .....	14
2. White Light Interferometry .....	18
2.1 Fringe Formation .....	19
2.2 Visualization of Fringes on Objects .....	21
2.3 Applying White Light Interferometry .....	23
2.4 True Color Images .....	24
3. System Concepts .....	25
3.1 First Concept – CMOS Thresholding .....	26
3.1.1 Implementation .....	26
3.1.2 Feasibility .....	29
3.1.3 Current State and Future of CMOS Technology .....	30
3.2 Second Concept – Two Sensor + DMD .....	31
3.2.1 Digital Micromirror Devices .....	32
3.2.2 Implementation .....	33
3.3 Third Concept – Three Camera + Mirror .....	34
3.4 Fourth Concept – Vision Sensors .....	38
3.4.1 Overview of Operation .....	38
3.4.2 Vision sensors .....	39
3.4.3 Detection of Fringes .....	40
3.4.3 Imaging Operation .....	41
3.4.4 Feasibility .....	41
4. System Design .....	42
4.1 Optical Design .....	42
4.1.1 Interferometric Microscope .....	43
4.1.2 Beamsplitters .....	45
4.2 Illumination .....	47
4.2.1 Köhler Illumination .....	47
4.2.2 Ring Light .....	49

4.3 Sensor Selection.....	50
4.3.1 Vision Sensor .....	50
4.3.2 Imaging Sensor .....	52
4.4 Final System Performance .....	53
4.4.1 Speed of Operation .....	54
5. Modeling .....	56
5.1 Coherence in FRED .....	56
5.2 Noise in Model: Gausslets .....	57
5.2.1 Bayer Filters .....	57
5.2.2 Amplitude Fluctuations.....	59
5.3 Artifacts from Discrete Wavelength Modeling.....	63
5.4 Roughness .....	67
5.5 Illumination Model .....	68
5.6 Imaging Arm.....	70
6. Conclusions.....	71
6.1 Future Work .....	73
6.2 Design Implications .....	74
References .....	75

# List of Figures

FIGURE 1. AIRY DISK AND ITS PROFILE. ....	11
FIGURE 2. TWO AIRY PATTERNS MOVING CLOSER TOGETHER UNTIL THEY ARE UNRESOLVED FROM ONE ANOTHER.....	12
FIGURE 3. DEPTH OF FOCUS FOR FAST AND SLOW SYSTEMS FOR THE SAME PIXEL SIZE. ....	12
FIGURE 4. PLENOPTIC CAMERA LAYOUT, INCLUDING MAIN CAMERA LENS AND MICROLENS ARRAY (LAM). ....	14
FIGURE 5. PROCESSING OF PLENOPTIC CAMERA IMAGE TO INCREASE DEPTH OF FOCUS (Ng). ....	15
FIGURE 6. THE (A) RECONSTRUCTION OF AN OBJECT FROM A (B) RADON TRANSFORM OF RAW DATA (NYGREN). ....	15
FIGURE 7. CONFOCAL MICROSCOPY AND ITS ABILITY TO IMAGE POINT-TO-POINT, IGNORING OUT-OF-FOCUS LIGHT. ....	17
FIGURE 8. LAYOUT OF A MICHELSON INTERFEROMETER.....	18
FIGURE 9. APPEARANCE OF FRINGES (WITH RESPECT TO OPD) FOR HIGH AND LOW COHERENCE SOURCES. ....	19
FIGURE 10. OVERLAP OF COSINE FRINGE PATTERNS FOR MANY INDIVIDUAL WAVELENGTHS, WHICH SUM TOGETHER TO PRODUCE WHITE LIGHT FRINGES. ....	20
FIGURE 11. IRRADIANCE ON SPECULAR, FLAT OBJECT WITH NO TILT, AT VARIOUS OFFSETS FROM PERFECT FOCUS. CONSTRUCTIVE (AT $\pm 0.05 \mu\text{M}$ ) AND DESTRUCTIVE ( $\pm 0.10 \mu\text{M}$ ) ARE VISIBLE, AS IS THE INCOHERENT LIGHT LEVEL ( $\pm 500 \mu\text{M}$ AND $\pm 1000 \mu\text{M}$ ). ....	21
FIGURE 12. IRRADIANCE ON A SPECULAR, FLAT, TILTED OBJECT, WHERE THE CENTER OF THE OBJECT IS AT PERFECT FOCUS. ....	22
FIGURE 13. IRRADIANCE MAPS MODELED IN SOFTWARE FOR DIFFERENT FOCUS POSITIONS WHEN MEASURING AN OBJECT WITH MULTIPLE FEATURES OF DIFFERENT HEIGHTS. NOTE ALL FEATURES HAVE THE SAME REFLECTIVITY WITH THE EXCEPTION OF THE RED SQUARE. ....	22
FIGURE 14. HIGH AND LOW CONTRAST FRINGES, HIGHLIGHTING THE DIFFERENCES IN AMPLITUDE OF IRRADIANCE FLUCTUATIONS. ....	23
FIGURE 15. THE APPEARANCE OF FRINGES IN DEPTH FOR AN OBJECT WITH VARYING HEIGHT AND REFLECTIVITY. THE LEFT AND MIDDLE FRINGES ARE THE SAME SHAPE (OBJECTS OF SAME REFLECTIVITY) BUT DIFFERENT HEIGHTS. THE RIGHT-MOST FRINGES SHOW THE OBJECT HAS DIFFERENT REFLECTIVITY (LOWER CONTRAST FRINGES) AND A DIFFERENT HEIGHT. ....	23
FIGURE 16. DESCRIPTION OF DEPTH RESOLUTION AND HOW IT MAY ROUND EXACT HEIGHTS OF AN OBJECT. ....	24
FIGURE 17. LENGTH OF FULL SCAN RANGE (500 MM) VS. THE NARROW RANGE FOR FRINGES TO OCCUR (A FEW MM). ....	26
FIGURE 18. TRANSFORMATION OF RAW FRINGE SIGNAL TO A THRESHOLDED SIGNAL, WHERE IRRADIANCE BELOW THE INCOHERENT LEVEL IS SET TO ZERO. ....	27
FIGURE 19. SCHEMATIC SHOWING A COMPARATOR PLUS MULTIPLEXER THRESHOLDING CIRCUIT. ....	27
FIGURE 20. TRANSFORMATION FROM RAW FRINGE SIGNAL TO TWO PROCESSED SIGNALS USING A DIFFERENTIAL AMPLIFIER AND A RECTIFIER, EITHER HALF-WAVE OR FULL-WAVE. ....	28
FIGURE 21. SCHEMATICS FOR THE DIFFERENTIAL AMPLIFIER (THE DIFFERENTIAL AMPLIFIER) AND HALF-WAVE (POWER DIODES AND RECTIFIERS) OR FULL-WAVE RECTIFIER (FULL WAVE RECTIFIER) CIRCUITS TO PROCESS THE FRINGES. ....	28
FIGURE 22. STACKED CMOS SENSOR WITH A LIGHT DETECTING, PIXEL LAYER ABOVE A LOGIC/COMPUTING LAYER (DEMOLDER). ....	30
FIGURE 23. SCHEMATIC OF CIRCUITRY AT THE PIXEL LEVEL FOR PRELIMINARY DESIGNS BY INIVATION, WHICH GREW INTO THEIR DAVIS SENSORS (LICHTSTEINER, POSCH AND DELBRUCK). ....	31
FIGURE 24. INDIVIDUAL MIRRORS OF A DMD TILTED IN BOTH THE "ON" AND "OFF" STATES, WHERE TILTING IS ALONG THE DIAGONAL OF THE MIRRORS (DOUGLASS).....	32

FIGURE 25. THE OPERATION OF A SINGLE MIRROR IN A DMD IN ITS "ON" AND "OFF" STATES WHEN USED FOR PROJECTION (GMUENDER).	32
FIGURE 26. LIGHT COMES FROM THE INTERFEROMETER TO THE DMD, WHERE IT IS DIRECTED TO EITHER THE SUMMING DETECTOR OR COMPARING DETECTOR, WHICH IS CONTROLLED AT THE PIXEL LEVEL.	33
FIGURE 27. SCHEMATIC OF JOHN SZE'S PROPOSED SYSTEM USING THREE CAMERAS AND A MIRROR TO SWITCH BETWEEN THE THREE CAMERAS.	34
FIGURE 28. EXAMPLE OF HOW THREE CAMERAS WILL READ A FRINGE PATTERN. IN THIS CASE, CAMERA 2 READS PEAKS, AND CAMERAS 1 AND 3 READ OFFSET FROM THE PEAKS BY $\lambda/3$ AND $2\lambda/3$ , RESPECTIVELY.	34
FIGURE 29. VARIANCE AND HOW IT CHANGES FOR DIFFERENT RELATIVE STARTING POSITIONS.	36
FIGURE 30. STRIATIONS APPEARING IN IMAGE OF A UNIFORM, TILTED OBJECT DUE TO THE CHANGES IN VARIANCE.	36
FIGURE 31. LOCATIONS OF MEASUREMENTS FOR A SINGLE CAMERA TO HAVE ITS MAXIMUM SIGNAL FROM FRINGES.	37
FIGURE 32. SCHEMATICS OF THE SCAN UP (INTERFEROMETER + VISION SENSOR) AND SCAN DOWN (IMAGING WITH SEPARATE DETECTOR) PROCESSES.	38
FIGURE 33. WORKFLOW DETAILING THE SYSTEM CONCEPT AND OPERATION OF THE INTERFEROMETER AND IMAGER TOGETHER.	39
FIGURE 34. SHOWCASING INAVATION SENSOR CAPABILITY TO COLLECT SYNCHRONOUS DATA (THREE BLACK AND WHITE FRAMES) AND PIXEL ASYNCHRONOUS EVENTS (INDIVIDUAL BLUE DOTS) (MUEGGLER).	39
FIGURE 35. DESCRIPTION OF FRINGES TRIGGERING EVENTS AND USING EVENTS TO FIND BEST FOCUS.	40
FIGURE 36. LAYOUT WITH INTERFEROMETER AND IMAGER COMBINED BY INCLUDING A SECOND BEAMSPLITTER.	41
FIGURE 37. LAYOUT OF WHITE LIGHT INTERFEROMETER, INCLUDING VARIOUS OBJECTIVE TYPES (YANG).	43
FIGURE 38. INTERFEROMETRIC MICROSCOPE LAYOUT WITH A LINNIK OBJECTIVE, HIGHLIGHTING LOCATIONS OF OBJECTIVES (BLUE) AND TUBE LENSES (RED) IN THE SYSTEM.	44
FIGURE 39. LAYOUT HIGHLIGHTING THE TWO BEAMSPLITTERS AND NECESSARY COMPENSATOR PLATE IN THE REFERENCE ARM.	46
FIGURE 40. GENERAL KÖHLER ILLUMINATION SCHEME.	48
FIGURE 41. LAYOUT OF A MICROSCOPE SYSTEM UTILIZING KÖHLER ILLUMINATION (ABRAMOWITZ).	48
FIGURE 42. SCHEMATIC OF IMPLEMENTATION OF KÖHLER ILLUMINATION INTO THE SYSTEM.	49
FIGURE 43. SYSTEM WITH KÖHLER ILLUMINATION AND ADDITIONAL RING LIGHT, USED TO INCREASE LIGHT FOR FINAL IMAGE.	49
FIGURE 44. VARIATIONS IN BRIGHTNESS OF COLLECTED IMAGE FROM DIFFERENT LIGHTING CONDITIONS (STEMMER IMAGING).	50
FIGURE 45. COMPARISON OF THE SENSOR SIZE FOR THE ORIGINAL REQUIREMENT AND VARIOUS VISION SENSORS.	51
FIGURE 46. FULL LAYOUT OF FINAL SYSTEM, DETAILING ALL COMPONENTS NECESSARY.	53
FIGURE 47. BREAKDOWN OF SPEED OF JAI GO-2400C-PGE OPERATION WITH TRIGGER FOR ROIS (JAI).	55
FIGURE 48. GAUSSIAN BEAM AND 5 OF THE 9 RAYS DEFINING IT. THE OTHER FOUR RAYS ARE DIVERGENCE AND WAIST RAYS INTO AND OUT OF THE PAGE (FRED APPLICATION NOTE: MODELING COHERENCE).	57
FIGURE 49. BAYER FILTER BUILT IN FRED (A) AND A REPRESENTATION OF A SINGLE BAYER FILTER "PIXEL" (B) WHICH CORRESPONDS TO FOUR DETECTOR PIXELS.	58

FIGURE 50. IRRADIANCE ON DETECTOR WITH BAYER FILTER FOR SMALL (A) AND LARGE (B) DETECTOR. A SMALL DETECTOR IS MODELED ACCURATELY, BUT THE LARGE DETECTOR IS AFFECTED BY ALIASING.....	58
FIGURE 51. TWO EXAMPLE PATTERNS SHOWING EFFECTS OF COHERENT MODELING WITH A BAYER FILTER. INCREASING RAY DENSITY NO LONGER IMPROVES UPON THESE PATTERNS.....	59
FIGURE 52. RIPPLES IN MODEL OF A "FLAT TOP" BEAM WHICH ARE CAUSED BY THE COMBINATION OF MANY GAUSSLETS.....	60
FIGURE 53. VISUALIZATION OF OVERLAP FACTOR, SHOWING BASE RAY AND GAUSSLET BEAM WAISTS. ....	60
FIGURE 54. THE EFFECT OF OVERLAP FACTOR IN THE MAGNITUDE OF THE FLUCTUATIONS FROM INDIVIDUAL GAUSSLETS. ....	61
FIGURE 55. FLUCTUATIONS IN IRRADIANCE OVER HALF THE FOV FOR VARYING RAY DENSITIES (RAY SEMI-APERTURE OF 51 TO 2001). NOTE THAT AS FLUCTUATION AMPLITUDE DECREASES, RINGING INCREASES ON THE EDGES OF THE APERTURE. ....	62
FIGURE 56. IRRADIANCE PLOTS SHOWING THE RIPPLE EFFECT AND HOW IT CHANGES WITH PROPAGATION DISTANCE. ....	62
FIGURE 57. SCHEMATIC OF TILTED OBJECT, TILTED BY ANGLE TO CAPTURE ALL POSSIBLE HEIGHTS OF AN OBJECT IN A 500 $\mu$ M DEPTH SCAN. ....	63
FIGURE 58. ARTIFACT IN THE FORM OF RESURGENCE IN FRINGE STRUCTURE IN FRED DUE TO USE OF DISCRETE WAVELENGTHS. ....	64
FIGURE 59. FRINGE RESURGENCE AS SEEN ON OUT-OF-FOCUS FEATURES OF THE OBJECT AND ON THE TOP AND BOTTOM EDGES OF THE FIELD. ....	64
FIGURE 60. FRINGES FOR 31 DISCRETE WAVELENGTHS, WITH VARIOUS RANGES FOR RANDOM OFFSETS OF WAVELENGTHS. ....	65
FIGURE 61. FRINGES FROM (A) LINEARLY AND (B) LOGARITHMICALLY SPACED WAVELENGTHS. ....	65
FIGURE 62. FRINGE STRUCTURE AND HOW IT VARIES WITH WAVELENGTH RANGE. ....	66
FIGURE 63. FRINGE RESURGENCE FOR COMBINATIONS OF GAUSSIAN AMPLITUDE, UNIFORM AMPLITUDE, LINEAR SPACING, AND LOGARITHMIC SPACING.....	66
FIGURE 64. MODELING OPTICALLY (B) SPECULAR AND (C) ROUGH SURFACES IN FRED AS THE SYSTEM MUST WORK FOR A (A) GENERAL OBJECT.....	67
FIGURE 65. FRINGES FROM VARIOUS SEMICONDUCTOR CHIP SURFACES, MEASURED WITH A ZYGO WLI. ....	68
FIGURE 66. LAYOUT OF SYSTEM MODELED IN FRED. ....	68
FIGURE 67. FRED MODELING RESULTS FOR ILLUMINATION ON REFERENCE MIRROR (A) AND TEST OBJECT (B).....	69
FIGURE 68. RELATIVE DIFFERENCE IN ILLUMINATION BETWEEN THE REFERENCE MIRROR AND TEST OBJECT.....	70
FIGURE 69. SAMPLE OBJECT IMAGED WITH FRED MODEL. 3x3 SQUARE PATTERN AT BEST FOCUS LOCATION. ....	70
FIGURE 70. AN OBJECT WITH THREE "WIRES" OF THICKNESS (1) 7.5 MM, (2) 5 MM, AND (3) 2.5 MM. ALL ARE IMAGED IN THE FRED MODEL WHEN AT BEST FOCUS. ....	71
FIGURE 71. LAYOUT OF DESIGNED SYSTEM WITH ALL COMPONENTS.....	73



## List of Tables

TABLE 1. REQUIREMENTS OF THE SYSTEM PRESENTED BY THE SPONSORING ORGANIZATION, AND THEIR LATER SHIFT TO NEW REQUIREMENTS ALLOWING SMALLER FOV TO ENABLE USE OF PROMISING SENSORS. ....	14
TABLE 2. QUALITATIVE DESCRIPTION OF THE FEASIBILITY OF THE THREE ON-CHIP PROCESSING METHODS, WHERE GREEN IS BEST AND RED IS WORST.....	29
TABLE 3. TRANSMISSION CHARACTERISTICS OF THE SECOND BEAMSPLITTER AND ITS EFFECTS ON THE RATIO OF LIGHT IN THE ARMS OF THE INTERFEROMETER AND IN THE AMOUNT OF LIGHT IN THE IMAGING ARM. ....	46
TABLE 4. COMPARISON OF SPEED AND EVENT OUTPUT TYPE FOR TOP TWO CONTENDERS: PROPHESEE ONBOARD AND INVATION DAVIS 346. ....	51
TABLE 5. COMPONENTS USED IN THE FINAL SYSTEM.....	53
TABLE 6. COMPARISON OF FINAL SYSTEM DESIGN PERFORMANCE COMPARED TO REQUIREMENTS. ....	72

## Abstract

An extended depth of field imaging system was developed for in-line inspection for the semiconductor industry. The system produces a single, two-dimensional, in-focus image of objects such as integrated circuits or other computer chip components. The system must be high resolution and have a large depth of field to capture small details at all heights of the object. The system must operate quickly since it is used on an assembly line and also must produce true color images, since color can be an indicator of issues with an inspected object.

The system uses white light interferometry (WLI) as a means of finding focus across the field of view. The interferometer vertically scans through different focus positions. Fringes from WLI only occur near best focus, or zero optical path difference (OPD). A “vision sensor,” which detects changes in irradiance, outputs events when the irradiance changes for each pixel, which only occurs in a narrow depth fringe region. A depth map of the object is created from this information and informs a separate imaging sensor of best focus depth for individual pixels. The imaging sensor captures pixels at their best focus, and the scanning combines the individual pixel information to create the in focus, two-dimensional image.

This thesis details the development of the system concept, the design of the system, and modeling of white light interference in general as well as its use in modeling system performance. The system designed has a 0.6-mm x 0.8-mm field of view with 1.5- $\mu\text{m}$  object size mapping to an image sensor pixel. The system has a 500- $\mu\text{m}$  depth of focus with depth resolution of 3  $\mu\text{m}$  and creates true color images. The prototype of the system is expected to create the final image in around 1 second. The speed is limited by capabilities of current detectors rather than the method itself, and it is expected the system can become significantly faster in a few years, such that full image capture can occur in a time of 40 ms.

# 1. Introduction

In optical systems, there is an inherent tradeoff between lateral resolution and depth of focus. Lateral resolution reaches a fundamental limit based on the width of the point spread function (PSF) of the system, which can typically be described by an Airy disk,

$$\text{Airy Disk diameter} = 2.44\lambda(f/\#), \quad (1)$$

where  $\lambda$  is the wavelength and  $f/\#$  is the F-number of the optical system. The Airy disk is the pattern (Figure 1) that appears due to diffraction from the edge of a circular aperture. Its size is inversely proportional to aperture size. Faster optical systems (low  $f/\#$ ) produce narrower Airy disk patterns than slow optical systems (high  $f/\#$ ). The Airy disk is the image of a perfect, infinitely small point source object, and as such its size becomes the factor that limits the lateral resolution of the system.

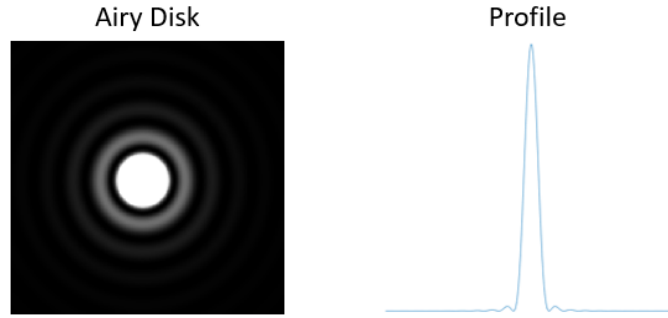


Figure 1. Airy disk and its profile.

Resolution is determined by overlap of two neighboring Airy disk patterns. As the patterns become closer to each other, they start to overlap and add together. The two Airy patterns are distinct and resolvable until they are close enough together that they cannot be resolved as two separate patterns (Figure 2), which occurs when they are separated by a distance  $d$  called the Rayleigh distance, given by the Rayleigh Criterion

$$d = \frac{0.61\lambda}{NA}, \quad (2)$$

where  $NA$  is the numerical aperture of the optical system, such that  $f/\# = 1/(2NA)$ . Narrower PSFs from faster systems are resolved from each other at closer spacings. The factor  $d$  describes the smallest resolvable feature, and to resolve small features, the  $NA$  must increase ( $f/\#$  decreases) to create a faster optical system. The Rayleigh Criterion specifically occurs when two PSFs overlap such that the peak of one PSF is located at the first zero or minimum of the other PSF, or when they are separated by the radius of the Airy disk.

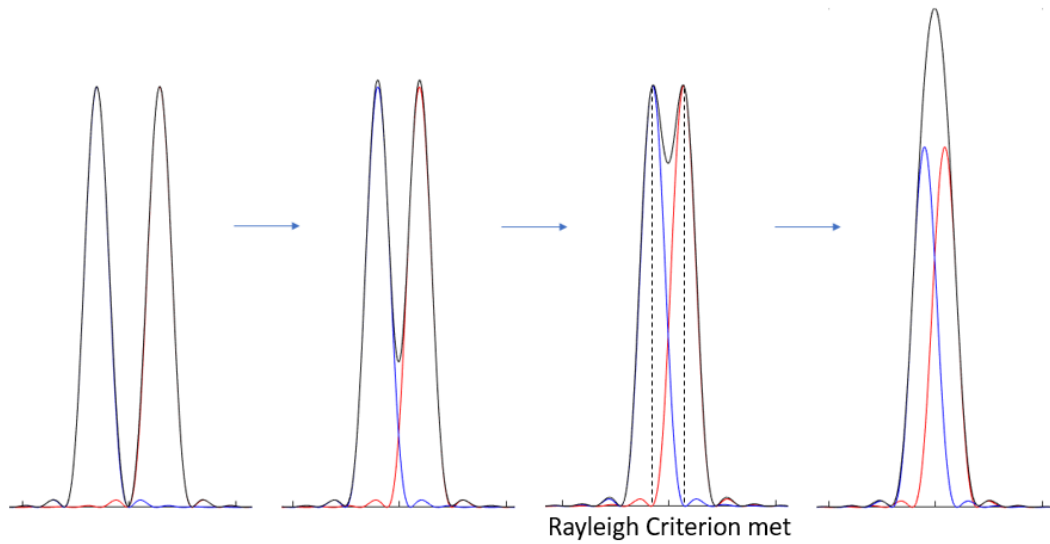


Figure 2. Two Airy patterns moving closer together until they are unresolved from one another.

However, when a system is made faster, its depth of focus (*DOF*) becomes shorter. The depth of focus describes the range of axial positions in an optical system around focus where the objects still appear to be in focus. There is an allowable blur in the system which does not affect the image produced, which allows the perfect focus to blur (defocus) until the spot size is the pixel size. If the pixel size remains constant, the depth of focus becomes shorter for fast systems and longer for slow systems (Figure 3).

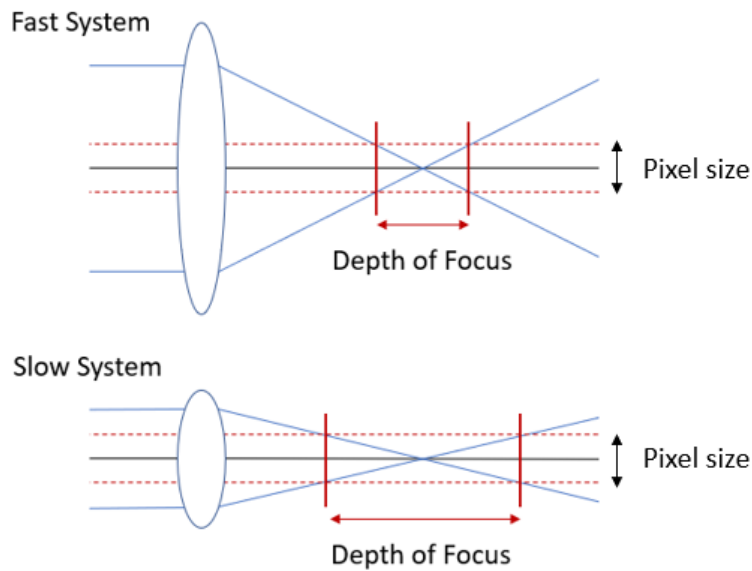


Figure 3. Depth of focus for fast and slow systems for the same pixel size.

Mathematically, the depth of focus is described by two equations, the first for objects beyond the hyperfocal distance and the second for closer objects,

$$DOF \text{ (beyond hyperfocal distance)} = 4\lambda(f/\#)^2 \text{ and} \quad (3)$$

$$DOF (otherwise) = 2(f/\#)_w \frac{m+1}{m^2}, \quad (4)$$

where  $(f/\#)_w$  is the working  $f/\#$  and  $m$  is the magnification. These confirm the direct dependence of  $DOF$  on  $f/\#$ , such that  $DOF$  increases with increasing  $f/\#$  (equivalent to decreasing  $NA$  and decreasing speed of an optical system). It is important to note the specific formulas to calculate  $DOF$  can change based on specifics of the system, however it inherently depends on  $f/\#$  in this way.

Lateral resolution and depth of focus are therefore opposing parameters: when resolution improves, depth of focus narrows. However, an optical system is desired which captures detail in three dimensions, so both high lateral resolution and large depth of focus are required. Systems that attempt to overcome the tradeoff between resolution and  $DOF$  are called extended  $DOF$  (EDOF) systems. EDOF systems enable small details to be captured at any height within an object, gathering information about the full object rather than a narrow depth (nearly a two-dimensional plane) within the three-dimensional object.

It is pertinent to note depth of focus is directly related to depth of field, and the two are used interchangeably in this thesis. Depth of focus describes the axial positions considered “in-focus” in image space while depth of field describes axial positions in focus in object space. Extending one parameter extends the other, so terminology is not greatly considered in discussion. The requirement of the system (Table 1) relates to depth of field, since EDOF system is based around inspecting an object with varying heights.

The desired output of the system is a single two-dimensional image that is in focus for the entire three-dimensional object being imaged. The system is applied to industrial, in-line inspection and is used to detect defects in semiconductor materials, such as small computer chips and integrated circuits. Inspection in an assembly line requires fast operation and imaging of small objects with very small features of varying heights. True color imaging is also necessary, since color of a component plays a role in its inspection.

The system can capture high resolution and large  $DOF$  images using a combination of optical and computational methods. The requirements of the system (Table 1) dictate many system characteristics, including object pixel size, which implies the lateral resolution of the system. The implied resolution and field of view also determine sensor size, which is 2,000 x 2,000 pixels.

The numerical aperture is found using the Rayleigh criterion (Equation 2), which requires the system to operate at 0.28  $NA$  or about  $f/1.7$ . This is a fast system which inherently has a short depth of field. The system is required to have a large depth of field (500  $\mu\text{m}$ ), such that an EDOF system must be used.

Table 1. Requirements of the system presented by the sponsoring organization, and their later shift to new requirements allowing smaller FOV to enable use of promising sensors.

Parameter	Original Requirement	Updated Requirement
Depth of Field ( $\mu\text{m}$ )	500	500
Field of View (FOV, mm x mm)	2.5 x 2.5	Not specified, can be less than 2.5 x 2.5
Depth Resolution ( $\mu\text{m}$ )	3	3
Object Pixel Size ( $\mu\text{m}$ )	1.25	1.25
Number of Images	1 – 3	1 – 3
Grabbing Time per FOV (ms)	< 40	< 40
Processing Time	Not specified	Not specified
Estimated System Size (mm x mm x mm)	180 x 25 x 50	180 x 25 x 50
Produce True Color Image	Yes	Yes

However, current EDOF technologies cannot meet these metrics. They often use many images and significant post-processing to create a fully in-focus image. Operating in this manner makes the systems easy to implement and use, but the systems become slow. The requirements demand innovation and improvement upon current EDOF techniques to create an efficient, fast, high resolution system that is well-suited to industrial inspection applications.

## 1.1 Types of EDOF systems

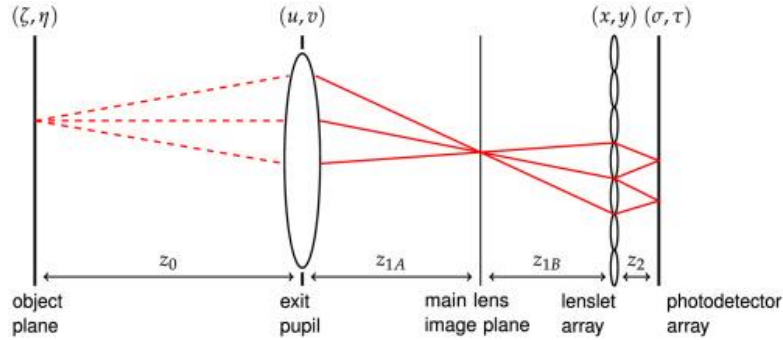


Figure 4. Plenoptic camera layout, including main camera lens and microlens array (Lam).

There are many types of extended depth of field systems which maintain high lateral resolution and increase system depth of focus. Systems using plenoptic cameras, image compression techniques, phase masking, and various scanning devices such as simultaneous temporal imagers, confocal microscopes, and interferometers all attempt to overcome the tradeoff between resolution and depth of focus. Systems can be limited by artifacts, require significant post-processing, operate slowly, or be unable to collect true color images. Each has its own features and drawbacks that are weighed to choose the general system type moving forward.



Figure 14: Refocusing after a single exposure of the light field camera. Top is the photo that would have resulted from a conventional camera, focused on the clasped fingers. The remaining images are photographs refocused at different depths: middle row is focused on first and second figures; last row is focused on third and last figures. Compare especially middle left and bottom right for full effective depth of field.

Figure 5. Processing of plenoptic camera image to increase depth of focus (Ng).

Plenoptic cameras use an array of microlenses implemented just before the image sensor, where each microlens captures its own sub-image of the scene (Figure 4). Each microlens collects a unique portion of the scene and by combining information from multiple microlenses, a 4D light field is determined. The light field can be used to extract depth information, since each microlens sees a slightly different portion of the scene from a slightly different angle. The raw image from the plenoptic camera can be processed to adjust focus (Figure 5),  $f/\#$  and the apparent angle of the camera. The image can be manipulated such that an  $f/4$  system can extend its depth of field to that of an  $f/22$  system. However, the system relies on significant post-processing to produce a single final image, making the system computationally intensive and slow. The microlens array also limits the resolution of the final image (Ng).

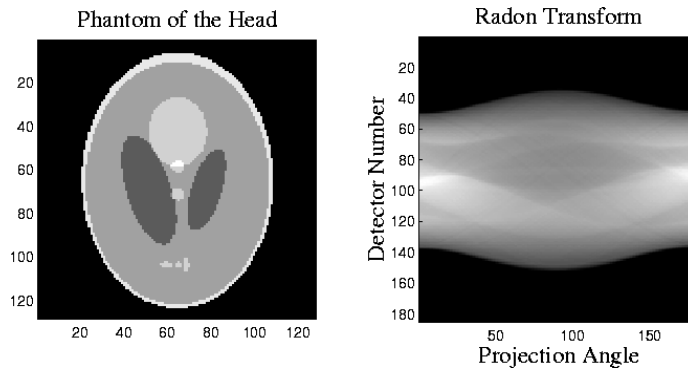


Figure 6. The (a) reconstruction of an object from a (b) Radon transform of raw data (Nygren).

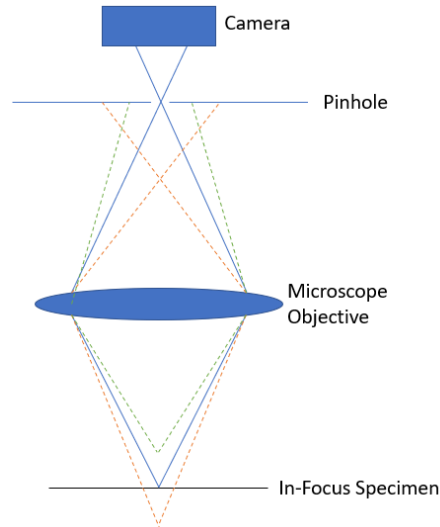
Compressed imaging is a faster method and is used in situations where data must be taken quickly. Less data is required than would be traditionally needed to create an image, but the final image is only achieved through significant post-processing. Such approaches are used frequently in medical imaging, where scans must be completed quickly as features can begin to move. Radon transforms are a typical method used in compressed imaging for this type of case (Figure 6). A Radon transform projects the object at specified angles and the information is collected at each angle. The angular slices are then back-projected on top of each other to produce an image. Such compressed imaging methods are not directly EDOF systems themselves, but instead speed up an imaging process to be used in an EDOF system. Typically, artifacts are still present in the final image, which is not ideal (Farber). In the case of Radon transforms, such artifacts include significant noise from the projection process and potential attenuation due to absorption of the light by the object.

Phase masks can be implemented in an aperture to change the point spread function (PSF) of an optical system in a predictable way. The changes to the PSF allow additional information about the scene to be extracted, but also degrade the PSF from its perfect form, sacrificing resolution in the process. The additional information also must be extracted from the raw scene, requiring post-processing of the image. Additionally, the masks can cause artifacts in the final image due to the strange PSF shapes that can occur (Castro).

Simultaneous temporal imaging methods scan vertically through different depth planes, collecting multiple 2D images that are processed and recombined to get the final high resolution image with a large depth of field. Such methods can be achieved with scanning objective lenses or deformable mirrors that change curvature. Simultaneous temporal imaging, however, operates too slowly. By scanning quickly enough to meet project requirements one ends up causing other problems mechanically, such as vibration. If the scans are done slowly in order to reduce such issues, multiple imaging channels are required, creating a more complex, expensive system with its own challenges, such as alignment. Additionally, there is significant computation required for this method in order to take all the images, analyze them, and recombine into the final image. (Duocastella).

Confocal microscopy is a method that utilizes pinholes to choose a specific depth in an object to image onto the detector. It is very similar to traditional microscopy, except these pinholes allow extraneous, out-of-focus light to be almost completely blocked from the detector, allowing a sharp, clear image of only the in-focus plane. In traditional microscopy, a certain field of view is visible to the detector, as defined by the optical system. In confocal microscopy, due to the limiting pinholes, only one point of the object can be seen by the detector at a time (Figure 7). To record a 2-dimensional slice through the object, the microscope scans across the required field of view. To collect depth information, there is another scanning mechanism to focus on planes of different depth positions.





*Figure 7. Confocal microscopy and its ability to image point-to-point, ignoring out-of-focus light.*

Point-by-point scanning is extremely time consuming, taking up to multiple seconds to scan a full 3D volume. There are a few variations on traditional confocal microscopy that make improvements, such as line scanning, arrays of confocal systems, and the use of Nipkow spinning disks. Implementing a Nipkow disk allows imaging through multiple pinholes at a time, which are formed in a spiral pattern on the disk, allowing faster scanning. Such systems have low illumination efficiency, but this can be improved drastically by introducing an array of microlenses above the Nipkow disk. The Nipkow disk can cause problems with crosstalk between different pinholes. The method improves greatly on confocal microscopy, taking a fraction of a second to scan the full 3D volume, however even the best confocal techniques appear to be about 10X too slow at the required resolution (Chong) (Jung).

White light interferometers scan vertically and produce fringes only when close to best focus, allowing focus to be found across the object quite easily. Additionally, the whole field of view can be seen at once, eliminating need for lateral scanning. Traditionally, white light interferometry (WLI) provides two-dimensional slices which are combined or otherwise processed to achieve a single in-focus image, which can be time consuming. White light interferometers can be used to get true color images since a white light (e.g., extended spectrum) is used as the source and color can be extracted from the fringes, however this process can slow down the operation further (Schmit, Novak and and Bui).

Each EDOF technique has its own unique challenges. Many techniques leave images with artifacts, require post-processing, or take significant time to gather an image. Despite its challenges, WLI is chosen as the method for the project. Typically, WLI is used when the depth information must be precisely determined. In the case of this system, depth is only needed to find best focus, which does not need much precision compared to typical WLI systems. Using WLI in this way is atypical and worth pursuing due to its novelty and the potential of the technology in this framework.

This thesis details the concepts of WLI to better understand how such a technique can be used in this application. There are also multiple system concepts presented, describing the progression of the design process with specific focus on feasibility. A fully detailed final system design is

described, indicating the operation, layout, and parts required to create such a system. Additionally, WLI in general and specifics of the system are modeled in software for better understanding and for confirmation of system performance.

## 2. White Light Interferometry

In typical interferometry, a high coherence source, such as a monochromatic laser (very narrow bandwidth), is used. This allows fringes to be visible for a large range of optical path differences (OPD) between the two arms of the interferometer. Coherent interferometric methods offer high quality results (sub-micron) for surfaces with small, slow changes, like curvatures of mirrors, and are often used to accurately measure such surfaces. However, this method is not well-suited to measure surfaces that have high slopes or sharp edges, as a  $2\pi$  phase ambiguity is inherent in the method. The fringes are cyclical, appearing the same when multiples of  $2\pi$  are added or subtracted from the OPD (e.g., OPD of 0,  $2\pi$ ,  $4\pi$ , etc. all appear the same), making the true OPD ambiguous.

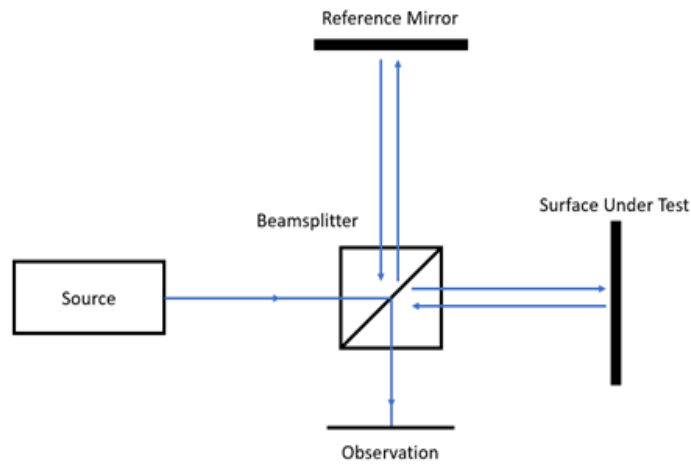


Figure 8. Layout of a Michelson interferometer.

In a double pass interferometer (Figure 8), the phase ambiguity limits measurable slopes to less than  $\pi$  (or  $1/2$  wavelength) per pixel (Wyant, 5.0 Direct Phase Measurement Interferometry). If the OPD changes by greater than  $\pi$  over a pixel, the OPD cannot be determined. When steep slopes or steps are measured, it is not possible to use the interferometric method with a high coherence source. This limits typical interferometers to “measurements of smooth, polished, homogenous surfaces” (Caber).

Typically, interferometers like this are only used to measure optical surfaces, and can only be used with visible wavelengths after grinding has occurred and once polishing has begun. An interferometer can use a long wavelength in the infrared to “obtain good contrast fringes when testing optically rough surfaces” (Kwon, Wyant and Hayslett) as rougher surfaces and steeper slopes are better measured with longer wavelengths. There is still a limit, though, in what slopes and roughness can be measured based on the wavelength used.

An interferometer can be used to measure step functions, higher slopes, and rough surfaces if a low coherence or incoherent source is used in the method called white light interferometry. The source with low temporal coherence (i.e., a high spectral bandwidth) significantly alters the fringe structure. Fringes are visible over a much smaller OPD range, when the OPD is very near zero. The incoherent source causes the fringe shape to change, where the maximum amplitude of the fringes occurs at zero OPD and the amplitude decreases until the incoherent level is reached, as shown in Figure 9. The incoherent level is indicated by relative irradiance of zero in the figure.

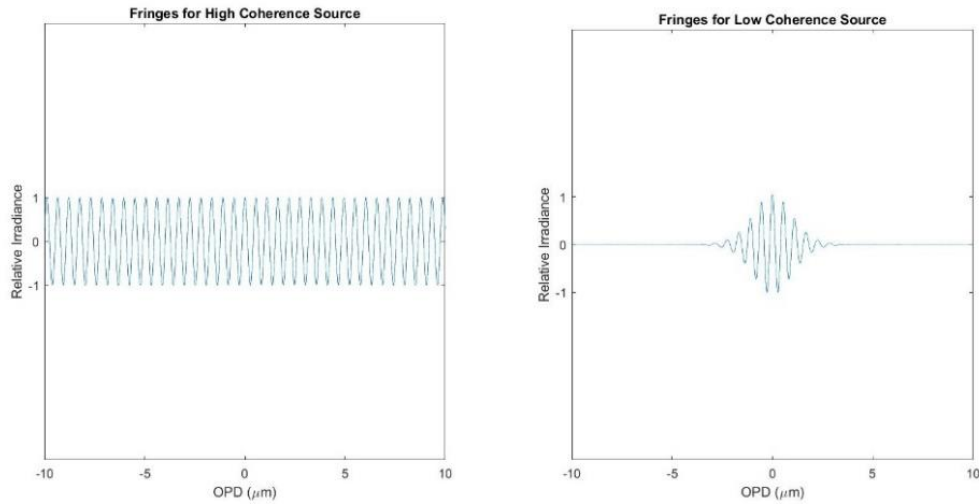


Figure 9. Appearance of fringes (with respect to OPD) for high and low coherence sources.

White light interferometry is limited at a certain point in how steep a slope and how rough a surface can be accurately measured. This is based on the resolution of the system and if “there are height differences within one resolution cell that exceed one-fourth of the wavelength of the light used,” (Pavliček) which introduces some ambiguity in what height is measured by the system for that resolution cell. The white light interferometry approach is not perfect but it is a significant improvement over other interferometry methods, especially in measuring rough surfaces with discontinuous or steep height variations.

## 2.1 Fringe Formation

The formation of the fringe structure in white light interferometry is best understood by considering the interference for each individual wavelength and simply summing the fringe irradiances to obtain the overall effect of the large bandwidth source. Each wavelength is highly coherent with itself and produces sinusoidal fringes such as in Figure 9. The fringe spacing is  $\lambda/2$  for a Michelson interferometer and the wavelength dependence causes each wavelength to have a unique fringe spacing. When all the individual fringe patterns of different spacings overlap (Figure 10), they sum together to create a peak at zero OPD, where all the patterns perfectly align. As OPD begins to change, the effect of the different fringe spacings for each wavelength begins to dominate, until fringe visibility becomes zero, as in Figure 9.

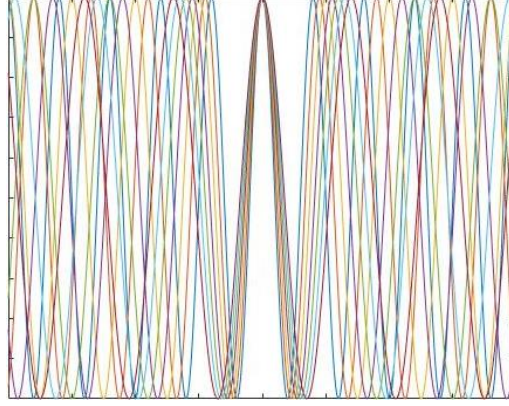


Figure 10. Overlap of cosine fringe patterns for many individual wavelengths, which sum together to produce white light fringes.

The specific OPD range where fringes are visible is highly variant and dependent on the light source used. Overall, the broader the bandwidth, the narrower the OPD range for visible fringes. Although this phenomenon is named white light interferometry, the color of the light need not appear white in color but simply must be broadband with low coherence, such that fringes only occur around zero OPD and there is no resurgence of the fringes at higher OPD.

The relative power in each wavelength affects the OPD range for fringe visibility ( $V$ ) and the relationship can be described mathematically. A source has a power spectrum which indicates the extent of its bandwidth and corresponding contributions from different frequencies. This is described by the spectral density, labeled  $G(\nu)$ , where  $\nu$  is frequency.  $G(\nu)$  is normalized to  $g(\nu)$ :

$$g(\nu) = \frac{G(\nu)}{\int_0^\infty G(\nu) d\nu} . \quad (5)$$

The power spectrum is transformed to have an area under the curve of 1, indicating the contribution of each spectral component, independent of the source power. This allows sources to be easily compared side-by-side to understand differences in spectral characteristics.

The temporal coherence visibility factor,  $\mu_{12}^\nu$ , is determined by  $g(\nu)$ ,

$$\mu_{12}^\nu = |\mathbf{F}_\tau[g(\nu)\text{step}(\nu)]|. \quad (6)$$

The value of  $\mu_{12}^\nu$  describes the extent the temporal coherence of a source affects visibility. When calculating  $\mu_{12}^\nu$ , the step function limits  $g(\nu)$  to only its positive frequency components. A Fourier transform is taken, producing  $\mu_{12}^\nu$  as a function of  $\tau$ , where

$$\tau = \frac{OPD}{c} , \quad (7)$$

where  $c$  is the speed of light. Low  $\mu_{12}^\nu$  values indicate fringes occur for smaller OPD ranges from zero OPD, as fringe visibility is directly proportional to  $\mu_{12}^\nu$ ,

$$V(\tau) = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2} \mu_{12}^\nu(\tau). \quad (8).$$

Here,  $I_1$  and  $I_2$  are the irradiance in the two arms of the interferometer. Because  $\mu_{12}^v$  is calculated primarily by using a Fourier transform, it becomes clear that wide bandwidth sources produce small temporal coherence visibility factors. Thus, understanding the spectral characteristics of a source is vital in understanding the fringes created by a white light interferometer.

## 2.2 Visualization of Fringes on Objects

It is important to understand how the fringe structure presents itself when inspecting a general object. There are multiple factors that affect the fringe appearance, including tilt, reflectivity, and feature height.

In the simplest case, a white light interferometer inspects an object under test which is perfectly flat, specular, and has zero tilt (normal to the optical axis). When the zero OPD condition is achieved, the entire plane produces constructive interference and the maximum irradiance at the peak of the fringe structure. As the white light interferometer scans in depth, the OPD changes and irradiance fluctuates. There are various areas of constructive and destructive interference within the fringes, and outside the OPD range for fringes, the irradiance is a constant, incoherent light level (Figure 11).

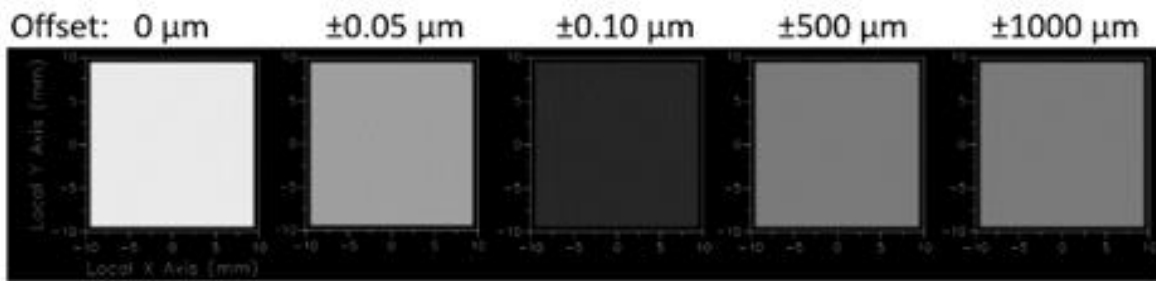


Figure 11. Irradiance on specular, flat object with no tilt, at various offsets from perfect focus. Constructive (at  $\pm 0.05 \mu\text{m}$ ) and destructive ( $\pm 0.10 \mu\text{m}$ ) are visible, as is the incoherent light level ( $\pm 500 \mu\text{m}$  and  $\pm 1000 \mu\text{m}$ ).

Tilt may be present in an object, and likely will be present to some amount in a real system. When an object is tilted, the zero OPD condition occurs for a line or region rather than the whole plane, resulting in a line or region with maximum irradiance. Moving to either side of the in-focus area, the out of focus areas of the object have an irradiance value corresponding to their specific OPD (Figure 12). For a flat surface with a prescribed tilt, a linear fringe pattern appears on the object. The width of the fringes depends on the amount the object is tilted, where less tilt creates wider fringes as there is less OPD variation. When scanning occurs for a tilted object, the linear fringe pattern on the object surface appears to move in the direction of tilt as a new area of the object becomes the location of best focus.

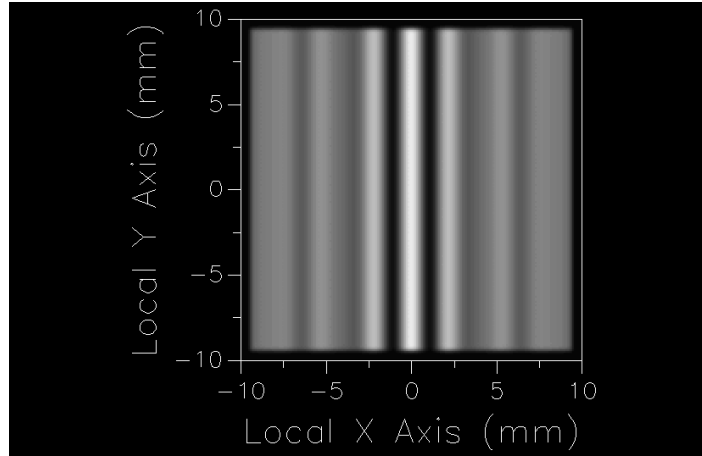


Figure 12. Irradiance on a specular, flat, tilted object, where the center of the object is at perfect focus.

When the object has features of varying heights, the fringes indicate when features are in focus. In Figure 13, there are four features in the object. If the background is the reference height of 0 mm, the feature heights are 0.1, 0.25, 0.4 and 1.0 mm. As the interferometer scans, features “light up” when they reach best focus, showing they are at the fringe pattern constructive interference peak. It is also clear when features are out of focus and out of the OPD range for fringes, as incoherent irradiance is maintained independent of depth. For example, the red square maintains the same irradiance (its incoherent level) for all example frames other than at its best focus of 0.25 mm (Figure 13).

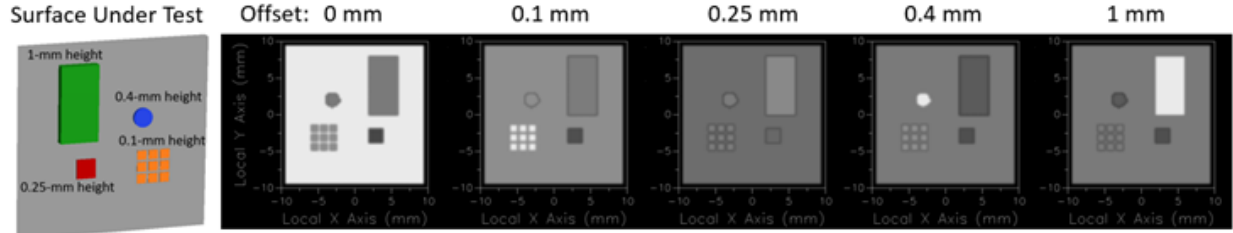


Figure 13. Irradiance maps modeled in software for different focus positions when measuring an object with multiple features of different heights. Note all features have the same reflectivity with the exception of the red square.

It is clear to see the peak irradiance value at best focus for the red square is not as “bright” as other features. The red square is lower reflectivity than other elements, showing up darker at both its incoherent level and in-focus, coherent position. The reflectivity of a feature affects the brightness of its fringes not only because its incoherent value is lowered, but also because the fringe contrast changes. Lower reflectivity creates smaller amplitude changes in the fringe pattern. The fringe contrast is reduced since the low reflectivity reduces the amount of light in the test arm for the object while the amount of light stays constant in the reference arm (Figure 14).

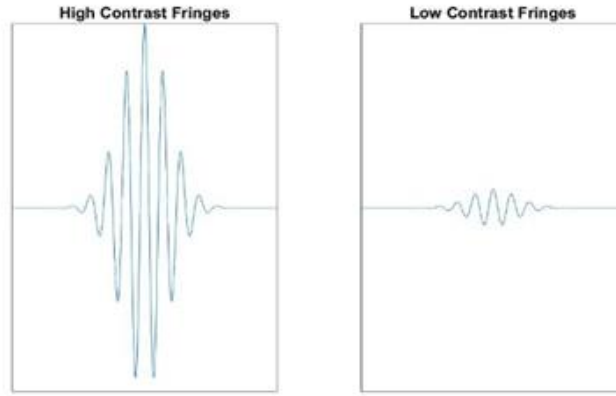


Figure 14. High and low contrast fringes, highlighting the differences in amplitude of irradiance fluctuations.

Fringe structure varies with location in the object due to height, overall tilt, feature tilts, and reflectivity. It is helpful to understand the fringe patterns relative to each pixel rather than globally. Each pixel has unique height and reflectivity characteristics and creates its own fringe pattern with certain contrast, which has peak irradiance value at a specific location (corresponding to the height of that pixel), as visualized by Figure 15.

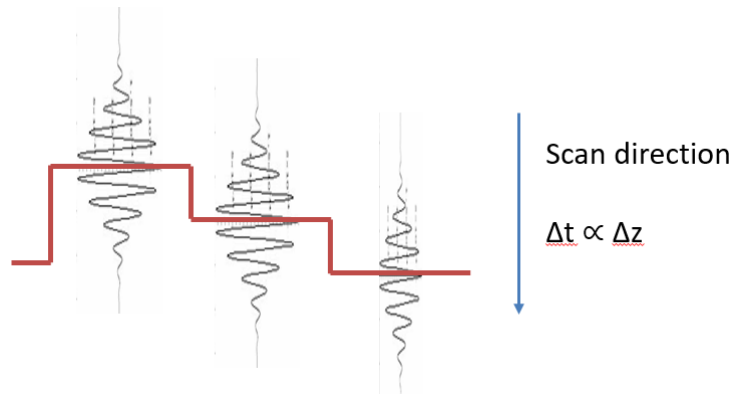


Figure 15. The appearance of fringes in depth for an object with varying height and reflectivity. The left and middle fringes are the same shape (objects of same reflectivity) but different heights. The right-most fringes show the object has different reflectivity (lower contrast fringes) and a different height.

Since the fringes of each pixel are independent, it is possible to measure nearly any object using a white light interferometer, making it ideal for this system. There is difficulty in measuring some objects, such as those with high slopes, but this is limited by the sensor or optical system and not the method itself. It also may be difficult to measure objects with low reflectivity due to low fringe contrast, but it is typical in many optical systems to have difficulty with low light levels.

## 2.3 Applying White Light Interferometry

Typical white light interferometry requires significant signal processing to create a 3D profile of an object. Many methods have been used for various WLI systems, including “signal processing electronics to demodulate the [interferometer’s] signal in real time” and “digital postprocessing of stored signals” (de Groot). These techniques are involved, often requiring Fourier analysis, sampling kernels, empirical models of fringes, or other algorithms to obtain the best focus based



on irradiance peaks or centroiding calculations. Processing is repeated for each pixel (de Groot). These methods produce depth maps with fine height resolution, around 1-5 nm independent of the interferometric objective type used (Schmit, An introduction to non-contact surface metrology). Since the entire FOV is measured at once and scanning only occurs vertically, WLI systems operate quickly relative to other EDOF methods. Other EDOF techniques – such as confocal microscopy – may require lateral scanning as well, which significantly decreases speed.

Despite the relatively quick operation of a WLI system, it does not operate quickly enough in this described form to meet system requirements. The required system must operate in 40 ms, where a system as described above may take multiple seconds to complete the same operation (Caber) (Schmit, An introduction to non-contact surface metrology). However, there are other requirements that allow the performance of the WLI system to loosen such that it can be significantly sped up.

The biggest difference in typical WLI operation and the system requirements is in depth resolution, or the smallest height difference in the object that is resolved. The requirement of this system is loose at 3- $\mu\text{m}$  depth resolution compared to the nanometer scale capabilities of typical WLI systems. This allows for less precision and removes the need for the complicated and time consuming algorithms required to precisely find a focus. Instead, the identified location of the “peak” of the fringes can occur over a less precise depth range while still meeting requirements (Figure 16).

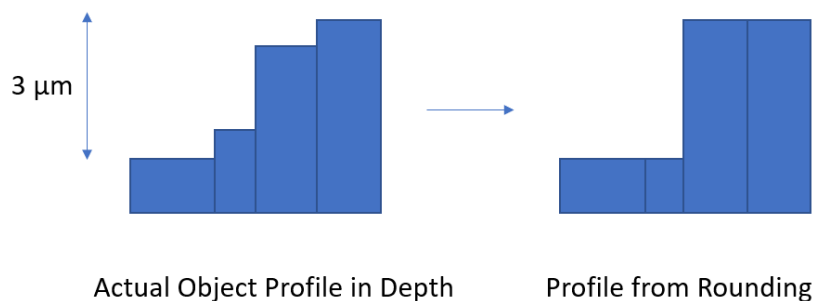


Figure 16. Description of depth resolution and how it may round exact heights of an object.

The required output of the system is a fully in-focus, two dimensional image rather than a three dimensional profile. The feature heights do not need to be recorded but instead are used to find focus. The precision in finding best focus affects the amount of blur in the 2D image rather than the precision of depth in a 3D map. There is a significant amount of blur (based on the inherent depth of field of the optical system) that can occur without impacting the final image, making the precision of depth less important. This added flexibility is utilized heavily in the design of the system.

## 2.4 True Color Images

It is possible to extract true color information from a WLI system, but doing so is complicated. The illumination from the reference arm and the fringe structure that forms can mask the true colors of the surface under test. The return from the reference arm can be much brighter than the



return from the surface under test, washing out color information. The fringes further cause an uncertainty in color since they fluctuate rapidly in spectral content and irradiance with changes in OPD.

Others have developed methods to create height maps and overlay true color measurements (Schmit, 2016). There are a few routes to take based on these methods. One method utilizes two scans. One scan gathers locations of best focus and a second samples pixels at best focus and gathers color information. This method requires two objectives and is slower since two scans are required. Another method uses side illumination with an incoherent source, allowing true color to be more visible. However, such a method can wash out the fringe structure and make best focus more difficult to determine (Schmit, Novak and Bui). It also creates a more complex system.

Frames can also be averaged around best focus locations to get a true color representation, which is simpler in its implementation since one objective is required and side illumination can potentially be omitted. At the zero OPD condition, the reflected light indicates the color of the object more clearly, since this condition shows the color reflected when the object is illuminated with white light. Fringes around best focus can be averaged to get a better real color image and remove the effect of the projected fringes (Schmit, Novak and Bui).

### **3. System Concepts**

Four system concepts were iteratively explored as challenges arose and different technologies inspired new ways of approaching the problem. The approaches differ from typical white light interferometry because of the different type of output required for this system. The system concepts are focused on transforming the fringe signal in a way to record the small OPD range where fringes occur. This OPD range for fringes is within the depth of focus of the optical system, meaning light can be measured over this OPD range and does not become significantly defocused or blurred over multiple pixels.

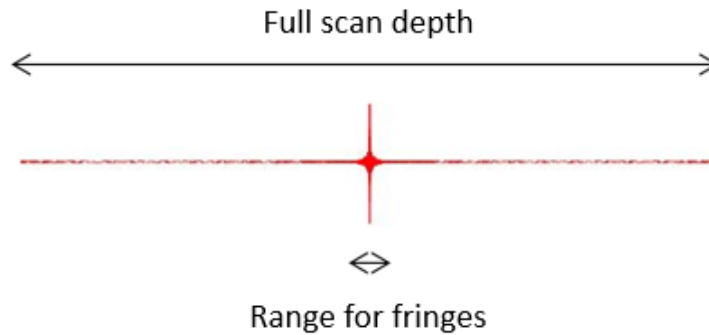
For better understanding, first consider the simplest, fastest method to record the fringe signal. This method has no data processing and uses a basic image sensor, simply summing as the scanning occurs over the full depth scan. In this way, each pixel records the incoherent light and the fringes. The incoherent light level is the signal for nearly the entire scan length, which is useless in the final image. Fringes – the useful information – occur for a very small region, and the constructive and destructive portions of the fringe structure cancel each other out, resulting in the incoherent level despite fringes being present. This method does not offer any way to collect a proper in-focus two-dimensional image.

From this thought experiment, it is clear a method is needed which performs some amount of processing on the fringes. However, this processing must be limited to achieve the high required speeds. The incoherent light level produces no information and is therefore ignored in the final image. Simple summation within the fringe region is just as useless, so the system must manipulate the fringes in some way to produce an image, which is the focus of the following system concepts.

### 3.1 First Concept – CMOS Thresholding

The simplest processing of the fringes is a threshold-type operation at the incoherent light level. An example threshold does not output for irradiances at or below the incoherent light level, and the system records only constructive interference.

Due to the speed requirements of the system, it is ideal to perform image processing steps “on-chip” rather than outputting full frames of data at each scan depth for postprocessing. In this system design, processing occurs on the sensor to manipulate the raw signal into its desired form. The sensor itself contains circuit elements to transform the data from each pixel without sending the signal off-chip to an external device such as a field programmable gate array (FPGA) or central processing unit (CPU) for processing.



*Figure 17. Length of full scan range (500  $\mu\text{m}$ ) vs. the narrow range for fringes to occur (a few  $\mu\text{m}$ ).*

On-chip processing avoids the bottleneck effect that occurs from serially transferring data off a sensor to another device, which often limits the speed of the system. Additionally, for most scan depths, a pixel is out-of-focus and simply at its incoherent light level which adds no information for the final image (Figure 17). If this information can be ignored and never outputted, rather than outputted and then thrown away, the system speed and efficiency increases significantly.

#### 3.1.1 Implementation

A threshold frame is recorded and saved on-chip. A somewhat out of focus, incoherent image is taken before the depth scan occurs. This image is stored and serves as threshold on a pixel-by-pixel basis; a global threshold cannot be set because the system must operate for any general object with features of varying brightness/reflectivity. The pixel irradiance is not expected to significantly change from the stored incoherent irradiance value until close to the zero OPD condition. Some amount of variation may occur in the incoherent irradiance region due to defocus and blur from neighboring pixels, so the thresholding may be set to an irradiance slightly greater than the collected reference image irradiance to avoid reading any pixels before reaching their near zero OPD condition.

There are three thresholding options that are possible to use for this system. One is a pure threshold scheme, as previously described, such that irradiance values at or below the incoherent level are ignored, and values above the incoherent level are maintained to create the signal (Figure 18).

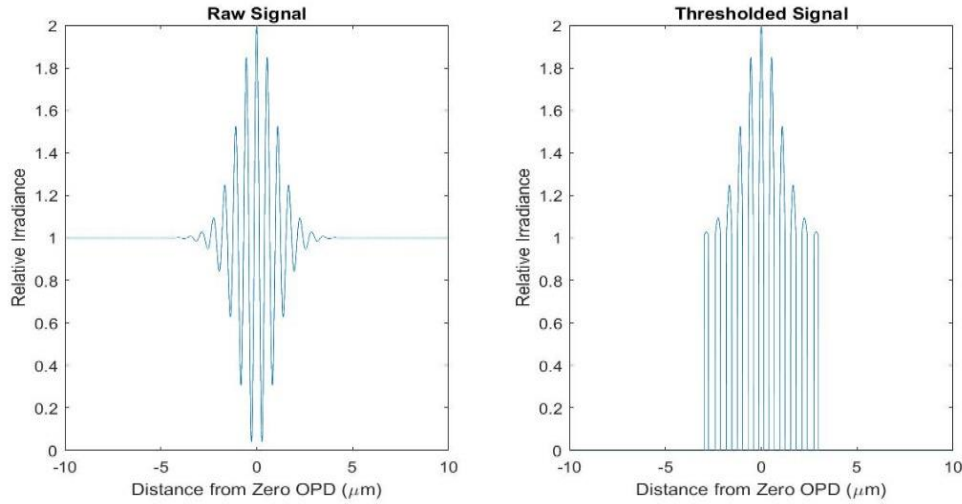


Figure 18. Transformation of raw fringe signal to a thresholded signal, where irradiance below the incoherent level is set to zero.

The thresholding scheme can be created with a simple circuit (Figure 19). A comparator is used to compare the input irradiance signal (e.g., voltage or other such term) from the pixel ( $V_{in}$ ) to the stored incoherent irradiance signal ( $V_{ref}$ ). The output ( $S$ ) of the comparator is high voltage if the irradiance is above the threshold value and is low voltage if the irradiance is at or below the threshold. The output of the comparator is used as the selector in a multiplexer. If the selector value is high, the first input ( $V_{in}$ ) is chosen as the output. If the selector value is low, the second input (0 V) is outputted. This circuit outputs  $V_{in}$  when irradiance is above the threshold and outputs nothing when  $V_{in}$  is below the threshold.

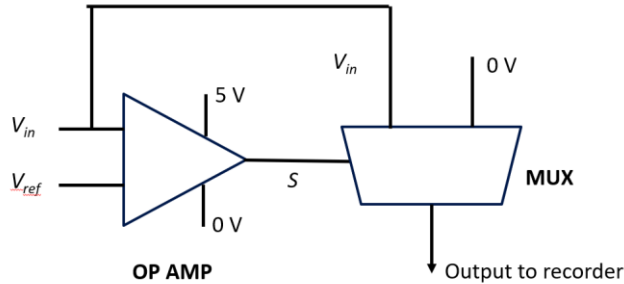


Figure 19. Schematic showing a comparator(using an operational amplifier) plus multiplexer thresholding circuit.

The other two options to process the fringe signal use a differential amplifier and a half-wave or full-wave rectifier. The outputted signals for these methods differ from the first method, since they transform the fringes in a different way. The half-wave rectifier method appears quite similar to the first method with the incoherent light level subtracted. The full-wave rectifier does this as well, but also flips the destructive interference to positive values to contribute to the signal (Figure 20).

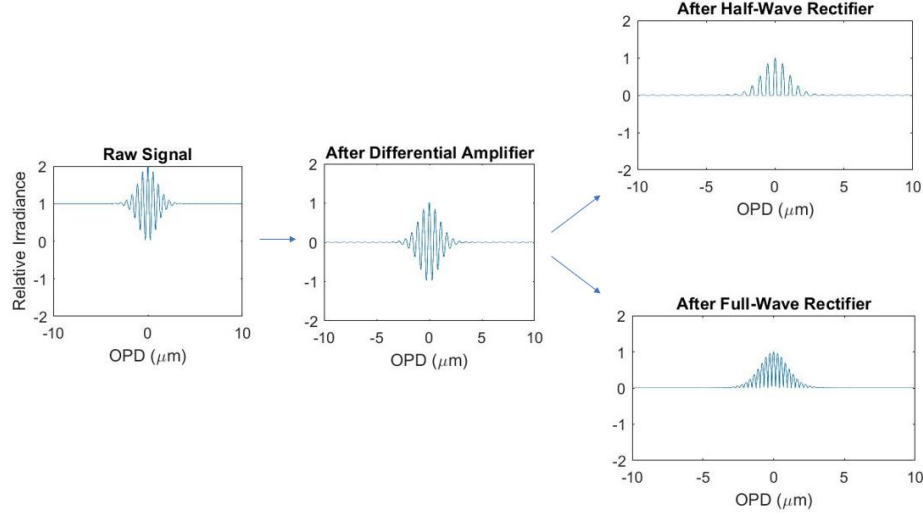


Figure 20. Transformation from raw fringe signal to two processed signals using a differential amplifier and a rectifier, either half-wave or full-wave.

The differential amplifier subtracts the stored incoherent irradiance from the pixel irradiance value. When the output of the differential amplifier is inputted to a half-wave rectifier, signals at or below zero are set to zero. Only the positive parts of the fringes are saved. When using the full-wave rectifier, the signal from the differential amplifier is transformed, outputting the absolute value of this signal rather than setting a threshold (Figure 21).

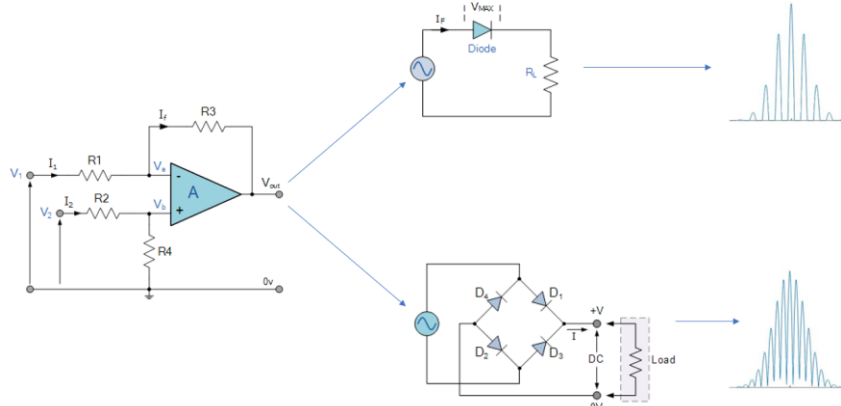


Figure 21. Schematics for the differential amplifier (The Differential Amplifier) and half-wave (Power Diodes and Rectifiers) or full-wave rectifier (Full Wave Rectifier) circuits to process the fringes.

Each circuit transforms the signals to record information from the white light fringes while ignoring incoherent light. If any of these circuits above were implemented on-chip, they allow the system to simply sum during the scanning, adding all the constructive interference (or constructive + “flipped” destructive interference for method 3) for each pixel over the entire scan depth. The peak of the constructive interference is perfect focus, and the signal on either side of focus is near perfect focus. The recorded fringe signal occurs for such a narrow OPD range that it is within the inherent depth of focus of the optical system and does not reduce system resolution.

### 3.1.2 Feasibility

Table 2. Qualitative description of the feasibility of the three on-chip processing methods, where green is best and red is worst.

Method	Irradiance	Contrast	Noise	Implementation
Comparator	Green	Yellow	Green	Green
Half-Wave	Red	Yellow	Green	Green
Full-Wave	Yellow	Green	Red	Red

The comparator method and half-wave rectifier method are inefficient in their use of the information the interference fringes provide. Instead of utilizing the signal provided by areas of destructive interference, these methods ignore them and set them to zero. The third method with the full-wave rectifier uses an absolute function value to include information from destructive interference in the processed signal, which allows for more signal from the fringes to be recorded. This is useful when smaller signals occur such as when contrast ratio is low (i.e., low reflectivity from a feature) (contrast column in Table 2).

However, the full-wave rectifier circuit does not set a threshold and any variations from the stored incoherent value are outputted. In this way, noise becomes a more significant issue in the system as it is defocused, since the threshold level cannot simply be increased to avoid recording such contributions as it can be in the other two methods. Additionally, any small fluctuations in irradiance are recorded rather than easily ignored as in the other two methods (noise column in Table 2).

The circuit of the third method is also more complicated and contains more elements, which is potentially problematic when implementing the circuit at the pixel level. The circuit diagrams provided are converted to transistors, and the third option is most requires the most transistors, taking up a larger area of the pixel and leaving less area for sensing material. This decreases efficiency of the sensor. As such, the first or second circuit method are most feasible for use in the system (implementation column in Table 2).

The first circuit method has a slight advantage over the second method as its relative irradiance values are higher since the incoherent light level is not subtracted. For low contrast fringes, this is beneficial. In the second (and third) circuit method, low contrast fringe signal may be interfered with by detector noise, such as dark current (irradiance column in Table 2).

There is potential to modify the processing circuits to apply them to individual rows/columns or small groups of pixels. This allows more area for sensing material and increases efficiency in space. The processing is not perfectly in parallel between all pixels, but can still limit the serial processing, especially when compared to outputting full frames for off-chip processing. However, this makes the storage and comparison to an incoherent level specific to each pixel complicated. There is also potential to use back-illuminated CMOS or stacked CMOS sensors to avoid the issue altogether, as the circuit components are placed behind light detecting components, removing losses in light collecting area due to circuit components.

It is possible to implement each of these schemes at the pixel level. Each required circuit element for each method can be represented by a combination of transistors and can be implemented in small integrated circuits. However, this specific thresholding type of processing is not available in commercial, off-the-shelf CMOS sensors. They can be designed and created, which is expected to occur in the future, but has not yet occurred. It is not reasonable to do such for this project, as CMOS sensor development is resource intensive, taking significant time and far greater funds than are available for this project (Ashok) (Roveda).

### 3.1.3 Current State and Future of CMOS Technology

It is predicted by the team that such a CMOS sensor with on-chip thresholding or other similar processing will be commercially available in the near future, perhaps in the next five years or so, since on-chip processing has become more common and a heavily researched area (Elouardi). Machine vision specifically has been a large driving factor of such sensor development due to the need for real-time and often complex computation.

One of the most recent contributors to this trend toward on-chip processing is CEA Leti, a French company, whose sensor “Retine” is a “3D stacked, backside illuminated, intelligent image sensor,” with stacked layers of photodiodes, processors, and memory. Processing does not yet occur for individual pixels but groups of 16x16 pixels, which is still quite an achievement and shows a trend toward processors on smaller and smaller pixel regions (Blackman). This sensor is more flexible than the hardware-based sensor described for the WLI application as the processors are programmable and able to operate independently. Such a sensor can be implemented in this WLI system because it is so general in its abilities and applications and aims to reduce output of extraneous data in the same way the WLI processing does.

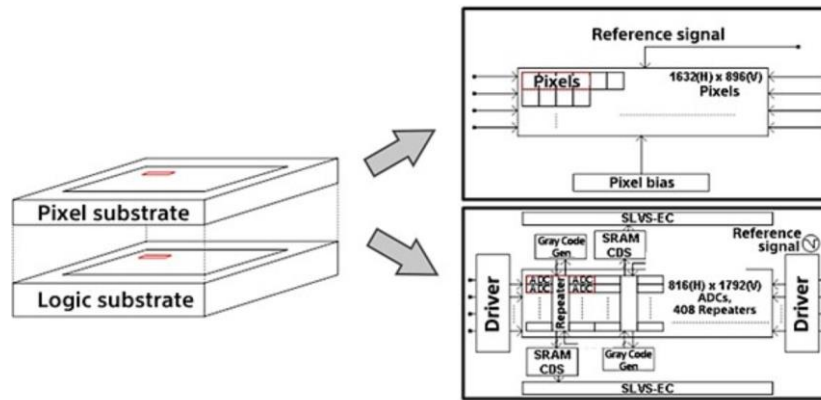


Figure 22. Stacked CMOS sensor with a light detecting, pixel layer above a logic/computing layer (Demolder).

Implementing the designed circuits in the way described is still possible in the future and it is not necessarily required to use processors as in CEA Leti’s sensor. There has been movement toward implementing circuitry components at the pixel level, which has been demonstrated and successfully used by Sony. Each pixel of a sensor is read out in parallel rather than in series by introducing an analog-to-digital converter (ADC) to each pixel. This allows for a global shutter which removes distortions in images caused by rolling shutters, and operates with lower power, higher speeds, lower noise, and higher dynamic range than previous global shutter sensors. The addition of the ADC at each pixel would normally take up too much of the area of the sensor, but

Sony has made the sensor stacked (Figure 22), with logic on the bottom and back-illuminated, photosensitive pixels on top (Demolder).

This technology is extremely promising as it has gone beyond the research level to be implemented in commercially available sensors. Implementing the additional circuitry required in the WLI system is not trivial, but is achievable, and this development by Sony shows a trend toward pixel-level circuitry capabilities.

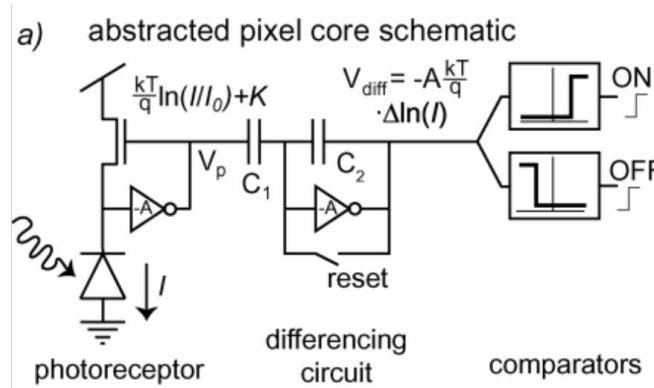


Figure 23. Schematic of circuitry at the pixel level for preliminary designs by Inivation, which grew into their DAVIS sensors (Lichtsteiner, Posch and Delbruck).

Even further, developments in “vision sensors” from companies such as Inivation and Prophesee have showcased pixel-level circuitry to create sensors that mimic human vision and detect differences (or changes in irradiance) rather than actual irradiance values. In the case of Inivation, the pixel contains “a fast logarithmic photoreceptor circuit, a differencing circuit that amplifies changes with high precision, and cheap two-transistor comparators” (Lichtsteiner, Posch and Delbruck). The circuitry implemented at each pixel (Figure 23) is similar in complexity to what is required for the WLI system. Since the writing of the paper by Lichtsteiner et al., Inivation has produced a commercially available vision sensor (as has Prophesee), meaning it is feasible and possible with current technology to implement such circuits at the pixel level.

These examples show the trend toward computing, processing, and transformation of signals on-chip before outputting information off a sensor, which is exactly the development required to make this circuit implementation possible in the coming years.

### 3.2 Second Concept – Two Sensor + DMD

The on-chip sensor requirements of the first concept are too stringent for current technology. The next concept attempts to alleviate some of the demands of the first technology by splitting the required functions between two detectors using a digital micromirror device (DMD). One detector is the “comparing detector” while one is “summing detector.” The light from the interferometer is directed toward one detector or the other by switching the angle of individual mirrors in the DMD.



### 3.2.1 Digital Micromirror Devices

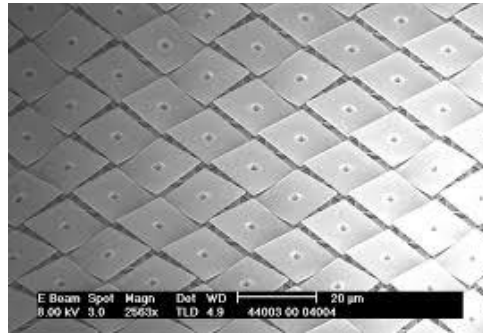


Figure 24. Individual mirrors of a DMD tilted in both the "on" and "off" states, where tilting is along the diagonal of the mirrors (Douglass).

DMDs appear somewhat like a detector but instead of an array of small light detecting pixels, a DMD is an array of small mirrors (Figure 24). Each mirror's tilt is controlled in a binary fashion, where there is an "off" and "on" state, with tilts of  $\pm 12$  degrees for most DMDs (Lee). DMDs can operate very quickly to switch between the two states, anywhere from 1.4 – 32.6 kHz (DLP Products: High speed visible - Products).

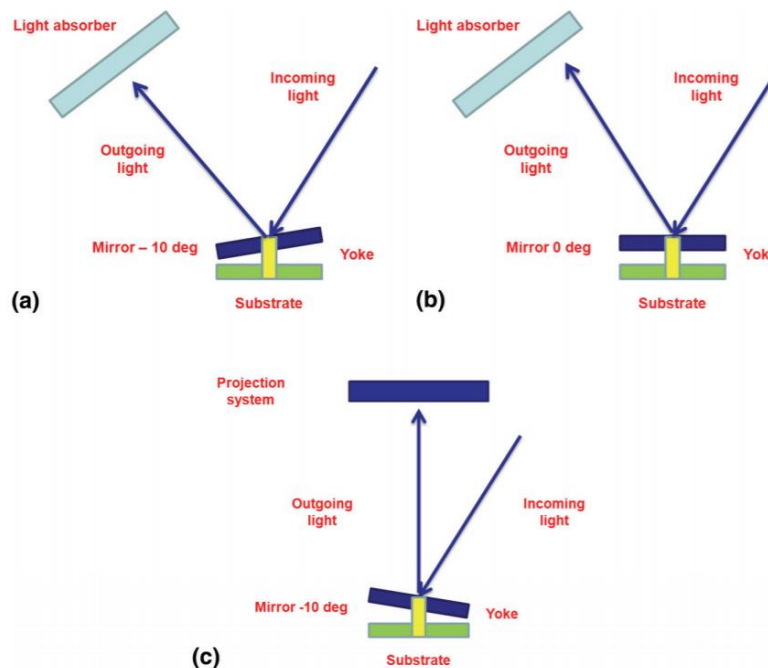


Figure 25. The operation of a single mirror in a DMD in its "on" and "off" states when used for projection (Gmuender).

DMDs are frequently used in projectors, where the "off" state points the mirror toward an absorber and the "on" state points the mirror into the projection system (Figure 25). Each mirror controls the light level of a single pixel in the projected image (Gmuender). Mirrors quickly switch between on and off states, and the brightness of the light projected from that mirror is based on the duty cycle of this switching.



### 3.2.2 Implementation

Instead of using a DMD to create pixels in a projection system, this concept relies on using a DMD as a spatial light modulator (SLM) to direct the light coming from the interferometer to the correct detector (Figure 26). One mirror of the DMD corresponds to a single pixel in each detector, mapping the pixels of one detector to the pixels of the other.

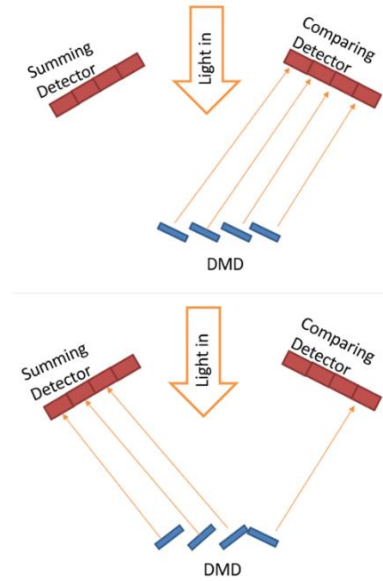


Figure 26. Light comes from the interferometer to the DMD, where it is directed to either the summing detector or comparing detector, which is controlled at the pixel level.

The “off” state points all the mirrors toward the comparing detector. This detector – like in concept one (Section 3.1) – compares the current signal on the detector to a stored threshold which is individual to each pixel. If the signal is above the threshold, the micromirror is turned “on” and points the light from that pixel toward the summing detector. The summing pixel is exposed to the light from the micromirror for a given exposure time, and after this, the micromirror returns to its “off” position.

This method is quite similar to the first concept but offloads the need for storage to a separate detector, rather than requiring it on the sensor that performs thresholding, reducing individual sensor load. The method is still head of its time in terms of the amount of circuitry that is required on the pixel level for thresholding on the comparing detector.

Additionally, this concept is challenging to align. Each micromirror of the DMD would have to be perfectly matched to its corresponding pixel in both the comparing and summing detectors, which is quite a difficult feat to achieve. Typically, researchers avoid having to align DMDs to full 2D image sensor arrays and instead align to single pixel sensors (Ouyang) or linear arrays. Some systems have successfully aligned DMDs to 2D sensors (Dumas) (Zhang), but only through extensive modeling and great care during actual alignment of the system which is difficult to achieve with one – let alone two – sensors.

### 3.3 Third Concept – Three Camera + Mirror

Another concept using multiple detectors was proposed by a member (John Sze) of the sponsoring organization. His concept uses three cameras and a single fast rotating mirror (not a DMD) to point the light outputted from the white light interferometer to three different cameras successively (Figure 27).

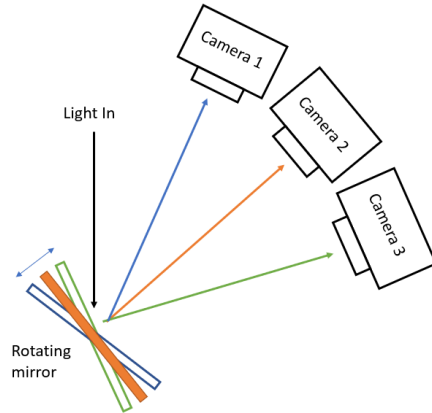


Figure 27. Schematic of John Sze's proposed system using three cameras and a mirror to switch between the three cameras

The mirror rotates to point at a different camera every  $\Lambda/3$  scanned, where  $\Lambda$  is the fringe spacing. The mirror points at each camera with a repetition rate of every  $\Lambda$  scanned such that each camera reads the same relative location on the fringes as scanning occurs. That is, if a camera reads the peak of a fringe, it will read the peak of all fringes, since they are spaced by  $\Lambda$ . The same holds true for valleys and every location on the fringes in between (Figure 28).

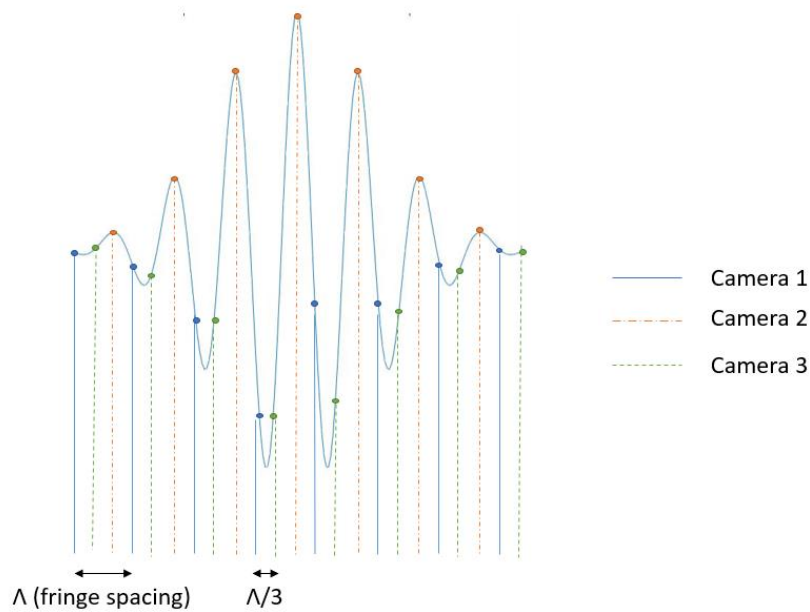


Figure 28. Example of how three cameras read a fringe pattern. In this case, camera 2 reads peaks, and cameras 1 and 3 read offset from the peaks by  $\Lambda/3$  and  $2\Lambda/3$ , respectively.

Each camera integrates its respective signal over the entire scan such that the incoherent level and fringe measurements as described are both recorded for each camera. The incoherent light is common to each camera, but the fringe signal is different for each camera, based on the location the camera measures on the fringe structure.

It is not known exactly where each camera begins measuring on each pixel's fringe structure since the pixel heights are random. If two cameras are used, they may produce the same signal depending on where their readings land on the fringe structure. With three cameras, it is guaranteed that there is some difference between the measurements of the three cameras, which indicates the presence of fringes.

The method measures variance between the three signals and use this variance as the fringe signal for each pixel. Variance is defined as

$$V = \frac{1}{N-1} \sum_{i=1}^N |A_i - \mu|^2, \quad (9)$$

where  $N$  is the number of measurements,  $A_i$  is the  $i^{\text{th}}$  measurement, and  $\mu$  is the mean. Variance measures how far away the samples are from the mean value. Variance indicates the coherent aspect of the light; the mean is the incoherent aspect, and any differences from the mean must come from the fringe structure.

For this concept to work, the variance must remain constant for the fringe pattern, independent of what points specifically on the fringes are read (given that the spacing between the measurements of each successive camera is maintained at  $\lambda/3$ ). This stipulation is required to create a proper image using variance.

Variance and its changes with “starting position” of the measurements on the fringes are analyzed. Starting position indicates where camera 1 takes a measurement on the fringe structure. If camera 1 reads at the very peak of the central fringe, this may be considered a relative starting position of zero. If camera 1 reads off-peak by  $+X$  OPD, then its relative starting position is  $+X$ .

There are only a small range of starting positions that must be analyzed. The cameras read points on the fringes separated by  $\lambda/3$ . If the starting position is  $\lambda/3$ , the same measurements are read as for the starting position of zero, just with a shift in which cameras take which measurements on the fringe. If camera 1 took a measurement at peaks at zero starting position, then camera 2 would take those exact same measurements at  $\lambda/3$  starting position. So, there are a limited number of starting positions that differ from each other, and only the range from 0 to  $\lambda/3$  need be considered to understand the changes in variance for all possible starting positions.

Unfortunately, over a range of 0 to  $\lambda/3$  relative starting position, the variance changes drastically (Figure 29). 5% deviation from the mean is indicated as well to show how large these variations are. This clearly shows such a method does not work measuring variance.

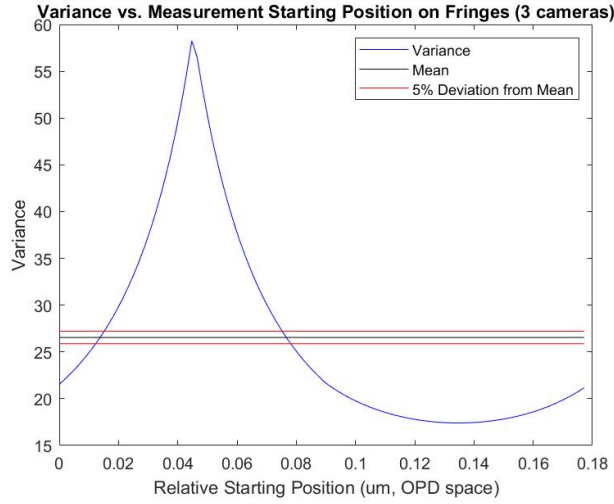


Figure 29. Variance and how it changes for different relative starting positions.

This is further confirmed by displaying how such differences in variation affect a two dimensional image produced by the system. If an object of constant reflectivity and color is tilted in the system, using the variance as the signal causes striations to appear in the image as an artifact of the method (Figure 30). The striations vary with tilt, with the spacing becoming narrower for larger tilt angles. It was then proposed by the sponsor that perhaps a different measure of sorts could be used instead of variance between the three cameras.

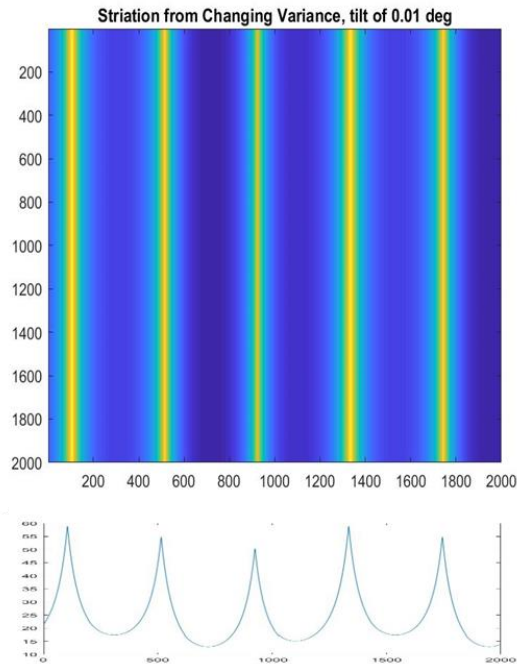


Figure 30. Striations appearing in image of a uniform, tilted object due to the changes in variance.

There is a more fundamental limit to this concept based on the signal-to-noise ratio (SNR). All the incoherent light collected during the long integration is noise, while the signal is the few points measured on the fringes by each camera. The cameras each read over a full scan depth of  $500\text{ }\mu\text{m}$

(Table 1) and each camera measures every fringe spacing  $\lambda$ . Each camera measures about 2,000 different depths. A single given pixel may have fringes for a region of a few microns so that around 10 of those 2,000 measurements are part of the signal rather than noise. Even for the maximum signal measured on the detector (Figure 31), the “noise-to-signal” ratio (opposite of SNR) is about 450. When different areas of the fringes are measured (due to different starting point locations), the ratio increases in favor of the noise. If the signal is 10%, 5%, or 1% of its maximum value, the ratio increases to 4,500; 9,000; and 45,000 respectively. At the limit of being able to detect this signal above the noise, it must be at least one “level” in the detector, where the number of levels is determined by bit depth.

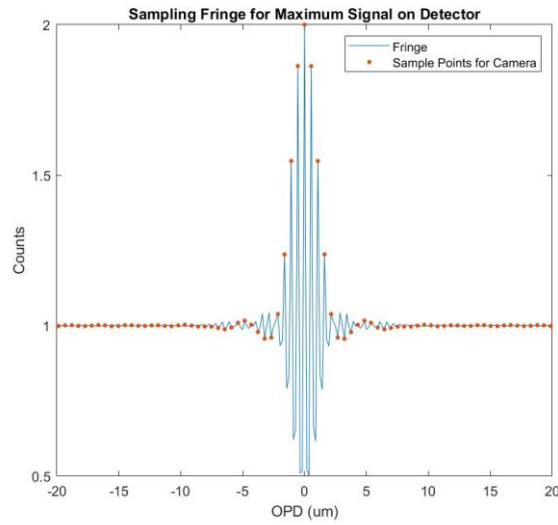


Figure 31. Locations of measurements for a single camera to have its maximum signal from fringes.

Such signals also must be able to be seen for low brightness objects as well; assume that the minimum brightness of the object is 10% of the maximum. With the addition of this condition, the number of levels that must be measured in the detector is 45,000; 90,000; and 450,000 levels respectively. The bit depth ( $n$ ) of the camera is calculated from this as

$$\text{Number of levels} = 2^n . \quad (10)$$

So, the required bit depth under these circumstances is 16-bit to 19-bit. However, the system does not detect the signal if the object is less than 10% the maximum brightness, the fringe contrast is  $< 1$  (due to non-specular surfaces), or there is any noise in the system.

To account for noise, the signal should be a minimum of 5 levels on the detector rather than 1. Additionally, to detect dimmer fringes from diffuse surfaces, another 5 levels are added, such that contrast down to 20% is visible. Bit depth increases to 20-bit to 24-bit, and these sensors are still quite limited in what objects can be measured.

Fundamentally, such a concept or any related concept that relies on summing over the full depth range and including incoherent light in measurements is ill advised. Sensors with such large bit-depths are scarcely available. Typically, 16-bit is the maximum bit depth available commercially,

and even these detectors are often limited by noise. Thus, further concepts must be explored for the WLI system.

### 3.4 Fourth Concept – Vision Sensors

Current sensors are limited in their on-chip processing capacity, and other developed concepts require careful alignment, are extremely sensitive, or require non-existent sensors. Still, it seems developments in current sensor technology are occurring and allowing for greater processing on-chip, but are not yet available on the scale (sensor size) required to maintain the large field of view of this system. Because using such sensors is promising, the requirements were updated (Table 1) to use smaller sensors, allowing exploration of interesting sensors.

#### 3.4.1 Overview of Operation

The method uses a vision sensor and a two-scan process. During “scan up,” the vision sensor (described in Section 3.1.3) is used as the sensor in the interferometer (Figure 32). The vision sensor detects differences (or changes in irradiance), making it well-suited to identifying when fringes occur. The interferometer and vision sensor are able to create a sort of 3D depth map of the object, as the vision sensor indicates best focus for each pixel.

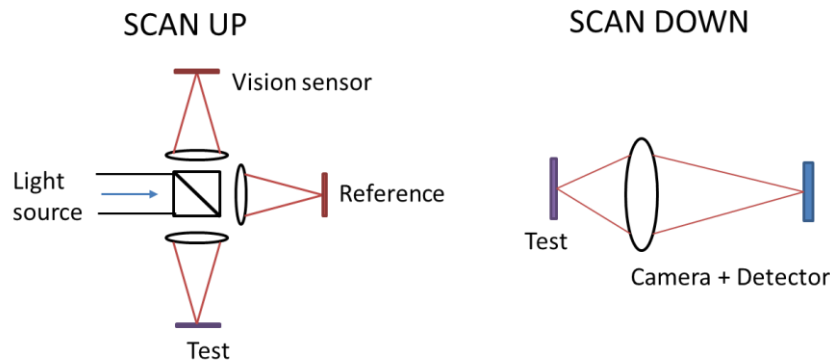


Figure 32. Schematics of the scan up (interferometer + vision sensor) and scan down (imaging with separate detector) processes.

During “scan down,” a more conventional camera system is used, separate from the interferometer, to record each pixel in focus. This process is informed by the depth map created in the “scan up” process. A full workflow of the system operation is detailed in Figure 33. To better understand system operation, vision sensors and how they can be used with fringes are further detailed, as are the imaging path and the integration of the two operations into one system.

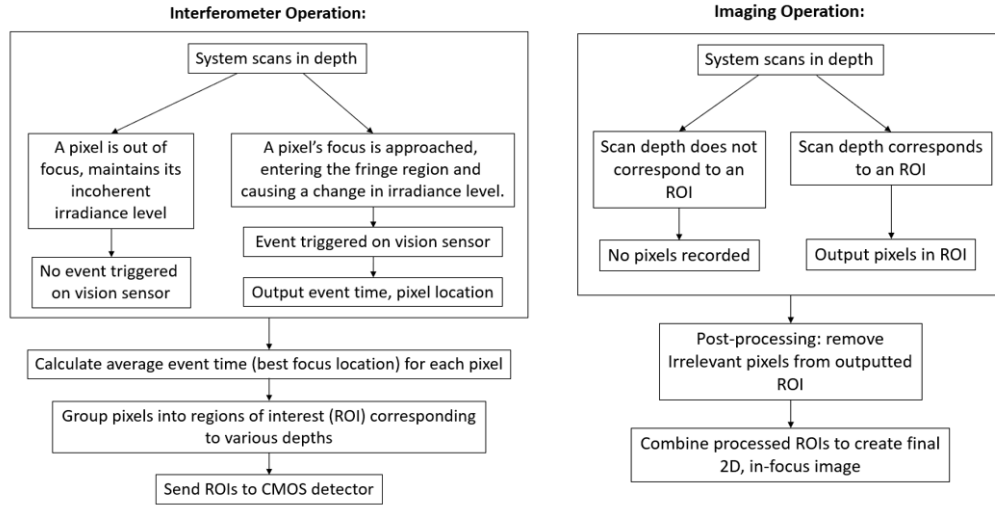


Figure 33. Workflow detailing the system concept and operation of the interferometer and imager together.

### 3.4.2 Vision sensors

Vision sensors operate differently than typical detectors. They are “difference detectors” that do not operate with fixed frame rates but “[rely] upon each pixel to only report what it sees when it senses a significant change in its field of view” (Prophesee). Vision sensors from companies Prophesee and Inivation are focused on limiting the output of extraneous data from a sensor. Their vision-like operation suppresses output of stable, unchanging backgrounds and outputs only pixels that are changing, inspired by the way human eyes are sensitive to movement but do not focus on unchanging features in a scene.

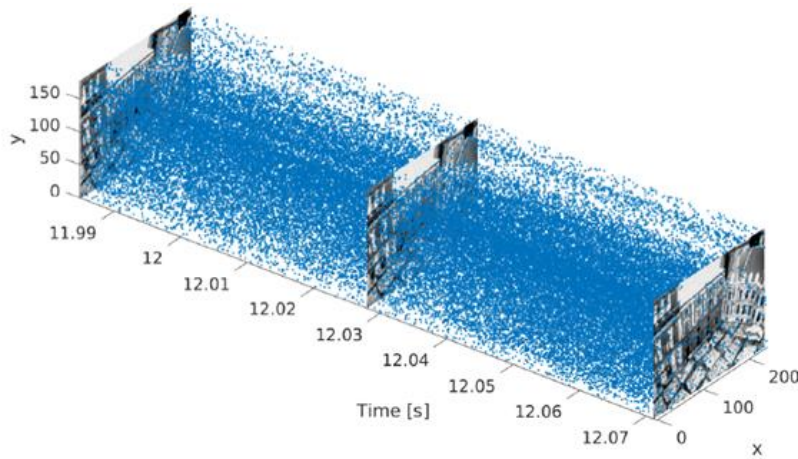


Figure 34. Showcasing Inivation sensor capability to collect synchronous data (three black and white frames) and pixel asynchronous events (individual blue dots) (Mueggler).

Inivation and Prophesee both require a change in irradiance on a pixel to output an event. Events are detected, in the case of Inivation’s DAVIS sensor, “in continuous time by specially-designed pixels” as “pixel-level, relative-brightness changes” (Mueggler). Events are outputted with pixel location, a timestamp, and an indication of whether the event was triggered from an increase (positive event) or decrease (negative event) in irradiance on the pixel. A series of outputted events can be seen in Figure 34. The level of change in irradiance that must occur for an event to be

detected is customizable in sensors from both companies, giving flexibility in the operation of the sensor.

Because the sensors do not output every pixel in a frame-like fashion, speed of outputting events is greatly increased. There is “low latency and high temporal resolution, both in the range of micro-seconds,” which means events can occur quickly and are outputted quickly, and another event can occur and be outputted quickly thereafter.

Each vision sensor is able to output binary events, showing a pixel either has an event or does not have an event. One detector from Prophesee is able to expand on this and produce “contrast detection events and exposure measurement events,” (Prophesee) which means it can provide true irradiance values. Inivation’s DVS sensors are not able to do so, but output binary events as well as full frames at a typical 30 fps rate, which for many applications allow irradiance of events to be interpolated (Mueggler).

In the case of this application, the fast speed of the binary events are crucial and most important, and a binary or true irradiance output is acceptable. The events provide location of best focus but no details on irradiance, so binary output is acceptable.

### 3.4.3 Detection of Fringes

A vision sensor is used as the detector for the interferometer. Outside of the small OPD range where fringes occur, the irradiance is nearly unchanging, constant at its incoherent light level. Thus, no events are outputted. When the fringes are approached, irradiance begins to change, and events start to be outputted. Events are outputted while irradiance changes, or when the depth is within the range for fringes for a pixel.

The timestamps outputted indicate the region where fringes occur, and the general location of focus. The median of the timestamps is when best focus occurred. The timestamps can be correlated to location of the scanning mechanism and therefore tell both time and depth of best focus (Figure 35).

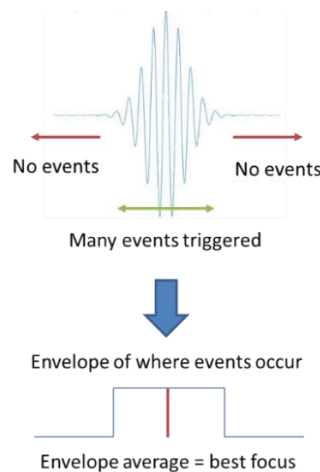


Figure 35. Description of fringes triggering events and using events to find best focus.



By performing this operation for each pixel, a depth map identifies the best focus locations for each pixel in the object. However, the depth map does not give true color information nor an image, but is used to inform a separate imaging system of the proper focus positions for each pixel.

### 3.4.3 Imaging Operation

The imaging path is incorporated into the interferometer system. An additional beamsplitter is inserted into the test arm such that the reflection from the beamsplitter points toward the imaging sensor, which collects the final 2D image (Figure 36).

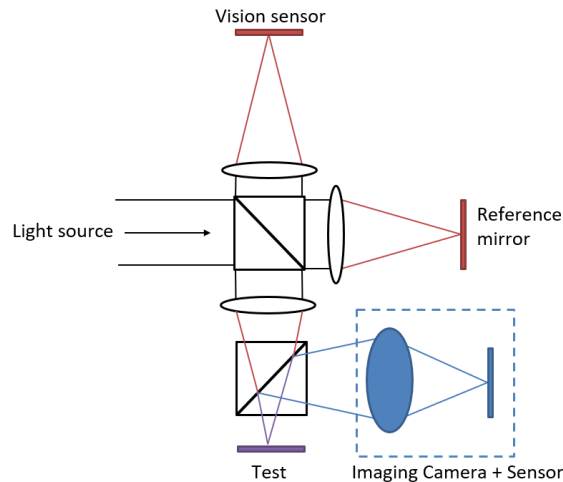


Figure 36. Layout with interferometer and imager combined by including a second beamsplitter.

The imaging sensor cannot be any general sensor but must have capabilities to use dynamic regions of interest (ROI) for fast operation. An ROI selects a specific area of the pixel array to output while ignoring other areas of the sensor. Outputting ROIs is faster than outputting full frames. The outputted depth map from the interferometer is grouped into ROIs. The perfect focus depth found for the depth map can be rounded to a reasonable degree (based on the depth of field and blur of the system) so that many pixels may make up one ROI corresponding to one best focus position. The imaging sensor receives this list of ROIs and their corresponding focus locations and cycles through them iteratively as each ROI comes into focus.

This system assumes there is no prior knowledge of an object and any object can be measured. So, ideally the ROI is any random combination of pixels that are in focus. Typically, ROIs are limited to rectangular shapes based on the CMOS architecture, which is based on readout in rows, columns, or grouped rectangular regions (Chouinard). The grouping of pixels to create ROIs becomes somewhat more complicated – the ROI does not translate exactly to pixels that are all in focus, but may contain extraneous pixels based on these ROI limitations. So, a bit of post-processing is required to select only the true in-focus pixels from each ROI.

### 3.4.4 Feasibility

This concept is most feasible amongst the four methods since it can be created with current technologies. Vision sensors exist and are able to perform quickly to output pixel information rather than full frames. The processing occurring between the raw output of the vision sensor and

the additional imaging sensor is limited. This step does simple averaging to find best focus positions for each pixel and then groups pixels into ROIs based on focus positions.

Some difficulties arise from only being able to record selected rectangular ROI, which involves some amount of optimization in the definition of the ROIs. It is best to have the fewest number of ROIs as possible for increased speed, but also to keep ROIs small for faster readout (Stemmer Imaging). So, an algorithm must be created for this method to best create ROIs, which is a drawback in the development of the system. But once a good optimization is created, it will operate quickly and allow the best and quickest combination of ROIs to be implemented.

Another challenge comes from alignment. Ideally, one pixel in the vision sensor should map to a single pixel of the imaging sensor. Alignment of two sensors at the pixel level can be quite challenging. This can be made easier by using an imaging sensor with a larger number of pixels such that a single pixel on the vision sensor maps to a 2x2 or larger region on the imaging sensor.

This concept is most promising because the challenges of the system can be overcome with current technology and the concept does not rely on technology which is still being developed in order to work. Additionally, applying a vision sensor in such a way is novel and an interesting concept to further develop. So, this system is chosen for the final design. Further characterization and design choices are described in order to characterize the full layout, function, and speed of the system.

## **4. System Design**

The system's optical design and components, illumination scheme, and sensor selection are detailed. The final system's commercial, off-the-shelf (COTS) components are listed and the system performance and requirements are compared.

### **4.1 Optical Design**

An interferometric microscope is chosen as the basis for the optical system. Interferometric microscopes are similar to typical microscopes, but differ in the type of objectives they use. There are three typical options for an interferometric microscope objective, which are what allow a microscope to be made into an interferometer. Michelson objectives typically have the lowest magnification. This interferometric objective uses a single microscope objective (common path to both interferometer arms) followed by a beamsplitter, creating a reference and test arm (Figure 37). Because the beamsplitter comes after the objective, the system must have a long working distance, which reduces *NA* and limits resolution capabilities. Michelson objectives, however, have the advantage of maintaining a large field of view.

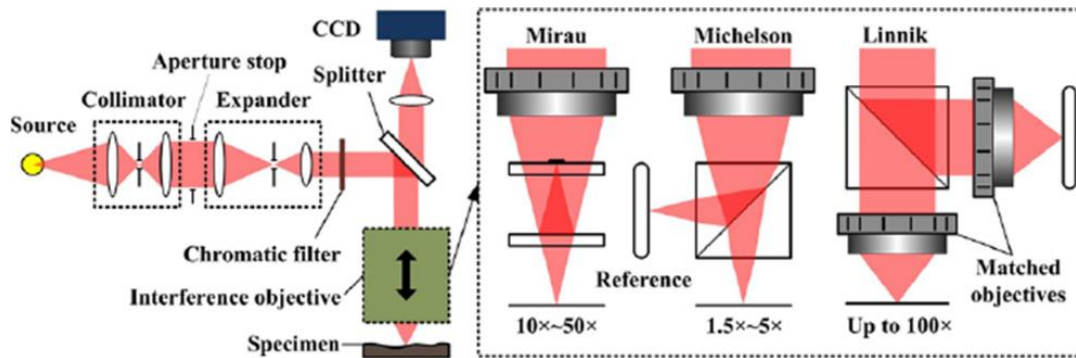


Figure 37. Layout of white light interferometer, including various objective types (Yang).

The Mirau objective has the next highest magnification. Like the Michelson, the Mirau objective utilizes a single microscope objective which is advantageous as it remains common path. It uses a beamsplitter plate and a reference mirror which is within the cone of light focusing on the object (Figure 37). This takes up less physical space than the Michelson components, allowing for shorter working distances and faster systems (higher  $NA$  and higher resolution). However, the reference mirror creates a central obscuration which has the potential to significantly reduce the amount of light through the system. The central obscuration size depends on magnification and gets larger with small magnifications. With large magnifications, there is not enough physical space to place both the beamsplitter and reference mirror in the focusing cone of light, meaning another objective type must be used.

The final objective type is the Linnik objective, which is unique in that it uses two microscope objectives. In this layout, the beamsplitter which creates the test and reference arms comes before any objectives. An objective is placed in both arms of the interferometer (Figure 37). These objectives must be well-matched optically and must not cause differences in OPD, dispersion, or aberration content between the two arms of the interferometer. Such a requirement can make this objective more expensive and difficult to create. However, it is much more versatile in its use, and its magnification,  $NA$ , and resolution are limited by microscope objective availability rather than inherent limitations of the interferometric objective design. Typically, the Linnik objective is used only when high magnification is required due to the added difficulty of using two microscope objectives, but it can also be used with lower magnification.

The choice of interferometric objective becomes a tradeoff between field of view and resolution. The Michelson objective has a large field of view and low resolution, the Linnik has a small field of view and high resolution, and the Mirau is somewhere in the middle, balancing between the other two objective's extremes (Wyant and Creath).

#### 4.1.1 Objective Selection

The original requirements of the system (Table 1) are difficult to achieve using a microscope system. The field of view (2.5 mm x 2.5 mm) is very large compared to the speed required ( $NA = 0.28$ ) to achieve the desired resolution (1.25  $\mu\text{m}$ ). For most optical systems, a 0.28  $NA$  requirement is quite limiting and cannot typically be achieved. Only certain types of systems can operate this fast, such as microscopes and microlithography systems. By a rule-of-thumb provided by optical design professor Dr. Jose Sasian, the approximated size of the optics of a system of this type should be about 10-12 times the full field of view diagonal (about 3.5 mm for this system), so in this case

the elements of the system are approximately 35 – 45 mm in diameter. Typical microscope objectives have an aperture typically around 5 mm and are around 12 mm long while microlithographic systems are about 600 mm in length. The requirements for the system place it between the sizes of these two systems, meaning the microscope has to be scaled up and become quite long with large lenses, or the microlithography system has to be scaled down, which requires many lenses and creates an expensive system (Sasian).

However, initial requirements were shifted to allow a proof-of-concept system to be created with vision sensors (Section 3.4). This shift makes the optical design of this system much simpler as well. The requirements shifted to maintain resolution while sacrificing a smaller field of view (Table 1), which is much more manageable and can be achieved with a traditional microscope system.

The choice of objective is not as simple as with a typical WLI system, since in this system a second beamsplitter must be inserted into the test arm to integrate the imaging portion of the system. This is difficult to achieve with a Michelson-type objective since working distance is limited and will not allow space for two beamsplitters. The Mirau-type objective uses a plate beamsplitter within the objective, which is difficult to modify to introduce another path. Additionally, all paths in the Mirau objective are coaxial, making it difficult to create an additional orthogonal path in order to insert the imaging sensor.

A Linnik objective is chosen due to its flexibility. It is easy to insert a second beamsplitter without issue. This presents challenges since the Linnik is the most difficult interferometric objective to work with; it is imperative that well-matched objectives are selected in order to have matching optical characteristics in the two paths of the interferometer. Despite the added difficulties, it is the best option because it is most easily modified to insert a second beamsplitter to create the imaging path.

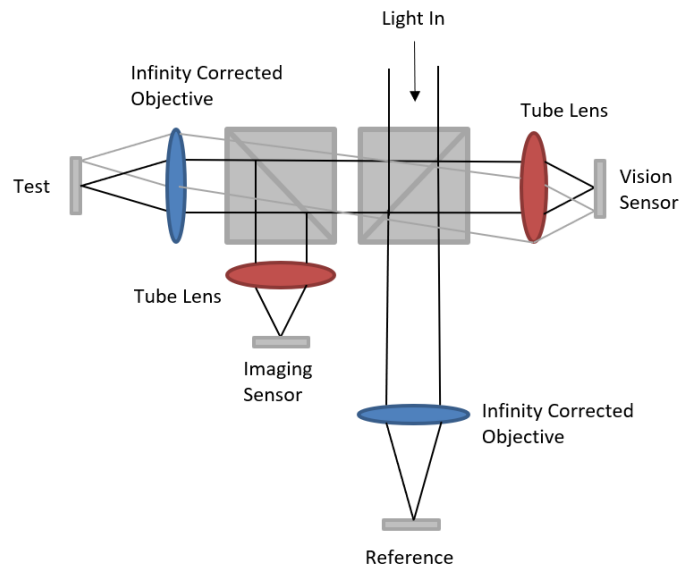


Figure 38. Interferometric microscope layout with a Linnik objective, highlighting locations of objectives (blue) and tube lenses (red) in the system.

Infinity corrected objectives were chosen for the system. These objectives collimate light rather than imaging to a point which allows for optical components – in this case, the two beamsplitters – to be easily placed in the system between the objective and the tube lens, which focuses the light onto the sensor (Edmund Optics) (Figure 38).

The choice of tube lens is limited – typically, the tube lens is implied by the objective selected. Objectives list a magnification, such as 10X, which is not specific to the objective but the objective-tube lens combination suggested by the manufacturer. So, once objectives are chosen, tube lenses easily follow to create a system with proper magnification. There is not room to stray from the objective-tube lens combination provided by a manufacturer. Optical designs or “black box files” which are provided for many optical components are rarely available for microscope objectives as manufacturers try to keep their designs as private as possible. Specific aberration content in an objective is not known and a tube lens or other lens system cannot be used in place of the manufacturer’s suggested tube lens. Doing so could introduce significant aberration, since the objective cannot be accurately characterized.

Magnification options are limited typically to 10X, 20X, 50X, and 100X (Edmund Optics). This has the potential to limit the system, since this magnification must be chosen based on the required magnification for the lateral resolution requirement and the vision sensor pixel size, which may not perfectly match one of the magnification options available. In this case, the objective is chosen based on a 15  $\mu\text{m}$  pixel size on the chosen Prophesee ONBOARD vision sensor (Section 4.3.1), where each pixel maps to an area of 1.25  $\mu\text{m}$  in the object. The ideal magnification is 12X, but the research team allows a 10X, 0.28 NA objective instead, such that resolution reduces slightly and the 15  $\mu\text{m}$  pixel maps to a 1.5  $\mu\text{m}$  region on the object instead.

#### *4.1.2 Beamsplitters*

The beamsplitters must be considered in the optical design because of their effects on OPD between the two interferometer arms. Also, the beamsplitter coatings determine the amount of light in each portion of the system, which is important to consider.

It is imperative in an interferometer to maintain the same optical properties in both arms, which is why there is concern with perfectly matching the two objectives. In the same way, there is concern in introducing the second beamsplitter in the test arm, as it produces different optical properties between the two arms. To match the two arms, a compensator plate must be included in the reference arm (Figure 39). The compensator plate is the same material as the beamsplitter and has the same dimensions so it affects the OPD and dispersion characteristics of the reference arm in the same way the test arm is affected by the beamsplitter.

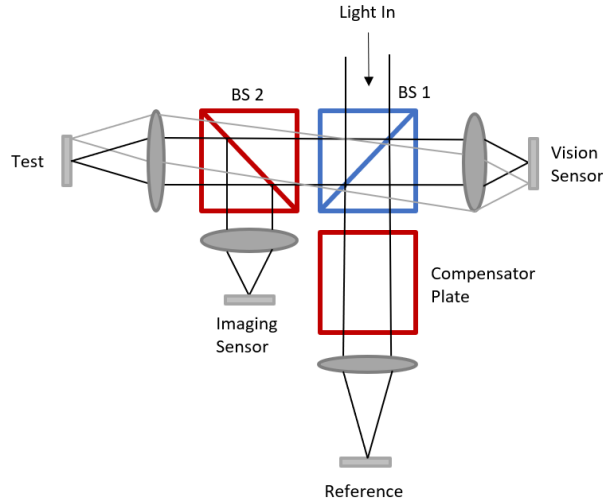


Figure 39. Layout highlighting the two beamsplitters and necessary compensator plate in the reference arm.

Additionally, the beamsplitter coatings must be carefully chosen to allow enough light to get to the imaging sensor while maintaining high fringe contrast. To best understand, let the transmission and reflection of beamsplitter 1 (BS 1) be  $T_1$  and  $R_1$  respectively, and of beamsplitter 2 (BS 2) be  $T_2$  and  $R_2$ . Each arm of the interferometer is affected by these parameters in different ways, which causes light levels to vary depending on location in the interferometer. This differs from a typical interferometer with a single beamsplitter, where the light level in each arm is simply  $R_1 \times T_1$ . In this system's case, the portion of the incoming light from each arm reaching the detector is:

$$\text{Interferometer Test Arm: } R_1 \times T_2 \times T_2 \times T_1, \quad (11)$$

$$\text{Interferometer Reference Arm: } T_1 \times R_1, \text{ and} \quad (12)$$

$$\text{Imaging Arm: } R_1 \times T_2 \times R_2. \quad (13)$$

The important considerations are the ratio between the two arms of the interferometer and the amount of light to the imaging arm. The value of the first beamsplitter has no effect on either of these considerations and is chosen to be 50/50 to allow for greatest amount of light in the interferometer. The second beamsplitter is not so easily determined. As the transmittance increases, the ratio of the power in the reference and test arm improves (ideally, the ratio is 1). However, the power in the imaging arm and therefore the amount of light available to collect the final image decreases (Table 3).

Table 3. Transmission characteristics of the second beamsplitter and its effects on the ratio of light in the arms of the interferometer and in the amount of light in the imaging arm.

Transmitted Power ( $T_2$ )	Reference Arm Power	Test Arm Power	Ratio	Imaging Arm Power
<b>50%</b>	0.25	0.0625	4	0.125
<b>70%</b>	0.25	0.1225	2.04	0.105
<b>90%</b>	0.25	0.2025	1.23	0.045

The exact choice for the second beamsplitter depends on the specific application. If the reflectivity of the test object and reference mirror are closely matched, it is possible to sacrifice some fringe contrast and allow lower transmitted power in order to get more light into the imaging arm. However, in this research application a general object with any level of reflectivity is assumed. If transmitted power is reduced, fringe contrast is reduced. Fringe contrast reduces further with lower reflectivity features in the test object, which can be detrimental to detecting fringes. So, high transmitted power is necessary.

Opposingly, more power is needed in the imaging arm in order to properly image the object. This is especially true with dim, low reflectivity, or highly scattering objects, where very little of the light inputted is reflected back into the system. So, the imaging arm power must be as high as possible in order to compensate for such issues, and so the transmitted power should be low.

A middle ground option is chosen for this WLI application. Table 3 presents the typical coating options available in COTS parts. The 70% transmitting, 30% reflecting option is chosen as it best achieves both conditions with a COTS component. In a more customizable system, the coating can be optimized based on the characteristics of typical objects measured to choose the best coating for the beamsplitter.

## 4.2 Illumination

The lighting of the system must uniformly illuminate the object and reference arm so characteristics of the fringes and final image come from object characteristics rather than lighting conditions. The lighting must be sufficient in brightness to detect fringes with the vision sensor as well as to image the object.

### 4.2.1 Köhler Illumination

In Köhler illumination, a collector lens collects light outputted from a source. A field lens images the source to the back focal plane of the condenser lens. The condenser lens illuminates the object with perfectly defocused light from the source, since it is imaged to its focal plane. The light source is imaged by the collector lens to an intermediate image plane, where an aperture stop is placed (Figure 40). For best illumination, the spacing is as described in Figure 40, which produces a uniformly illuminated area on the object (Mathew). Using such an illumination scheme produces uniform illumination on the object even when the light source may not output in a uniform way. Additionally, it avoids imaging the shape of the light source itself to the object, which is an issue with critical illumination is used (Arecchi, Messadi and Koshel).

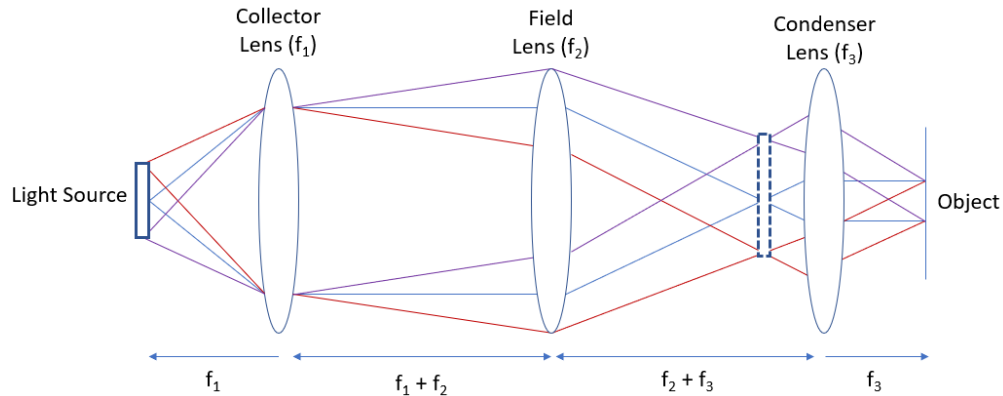


Figure 40. General Köhler illumination scheme.

When applied to microscopy with brightfield, reflected illumination, the condenser lens is replaced with the microscope objective. The objective already focuses on the object to image it and directs light to the object in the same manner. The source and collector and field lenses are necessary to complete the illumination scheme (Figure 41). Often in microscopes, aperture and field stops are inserted to adjust the illuminated field of view and the ray angles out of the object illuminating the specimen. This is not necessary in this system, as it is designed already for a specific field of view and with a specific  $NA$ . Additionally, two lenses for the collector and field lenses is the minimum, but more lenses can be used as necessary to create a proper illumination scheme.

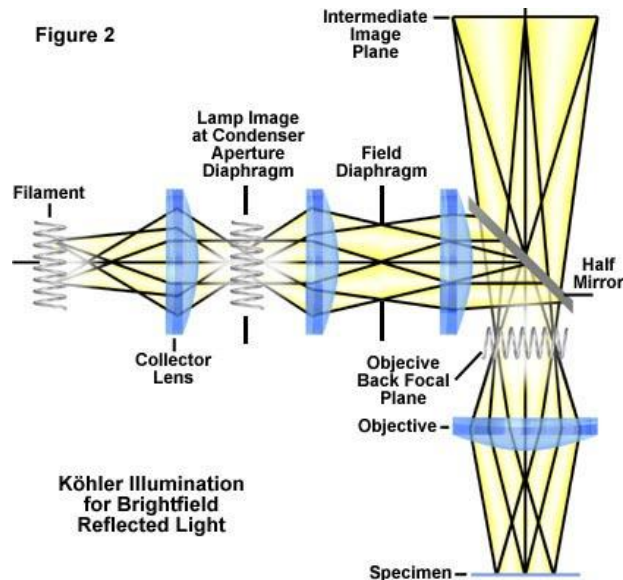


Figure 41 Layout of a microscope system utilizing Köhler illumination (Abramowitz).

Because of its uniformity and the relative ease of implementation, Köhler illumination is used in this system as the lighting scheme (Figure 42). It illuminates samples in an acceptable way and is flexible with what light source is used with it. It also allows for best true color and brightness representation for the object, since it illuminates all parts of the object evenly, so differences in brightness in the image come from variation in reflectivity over the object.



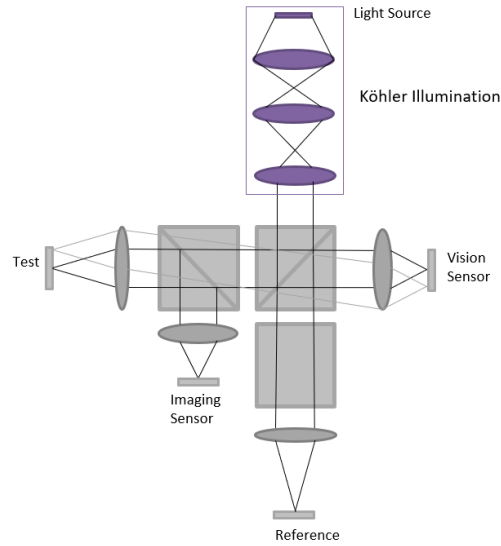


Figure 42. Schematic of implementation of Köhler illumination into the system.

#### 4.2.2 Ring Light

There is likely low lighting in the imaging arm of the system. It is ideal if the light for the final image can be increased. Such a thing can be achieved by introducing another illumination element in the form of a ring light surrounding the objective that faces the test object (Figure 43).

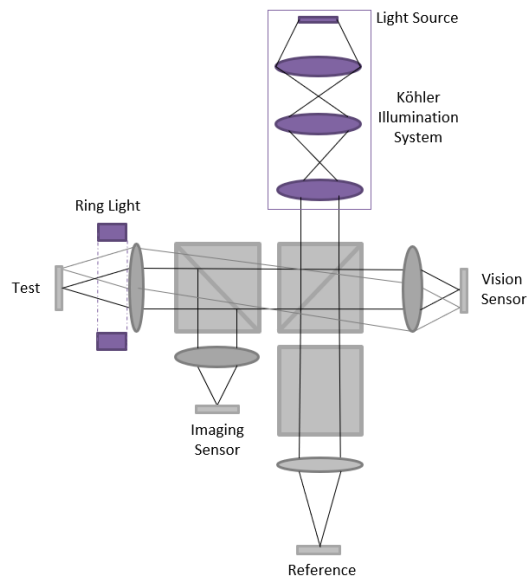


Figure 43. System with Köhler illumination and additional ring light, used to increase light for final image.

In system use, the ring light appears to flash; it is off when the vision sensor and interferometer are operating and turns on to illuminate the object when the image is taken using ROIs. When imaging, the object is illuminated by the Köhler illumination scheme and the ring light, greatly increasing the amount of light to the image sensor despite the issues introduced by BS 2 (Section 4.1.2). In fact, introducing a ring light allows further flexibility in the coating of BS 2. Perhaps a

90/10 coating can be used due to the additional amount of light available for the imaging operation, allowing a better ratio to be achieved between the arms of the interferometer.

However, introducing a ring light changes the illumination conditions between the interferometer and imaging operations. The Köhler illumination provides coaxial, uniform lighting while the ring light provides direct illumination from the side. Objects appear differently when illuminated with side rather than coaxial illumination. The ring light has the potential to create non-uniformity in the illumination of the object, which would appear in the final image. This could cause a feature of the object to appear brighter or dimmer depending on illumination rather than the reflectivity of the object, which creates ambiguity in the cause of variations of brightness in the image (Figure 44). It also introduces the possibility of shadows being cast from various features in the object.

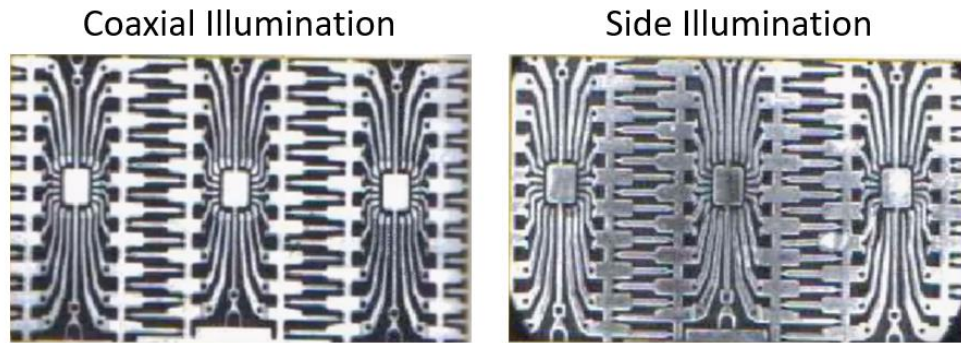


Figure 44. Variations in brightness of collected image from different lighting conditions (Stemmer Imaging).

A direct reflection of the ring light could also appear in a final image if a feature of the object is specular such that it reflects the ring light directly to the sensor. This artifact would be clearly seen in the image and does not accurately represent the object alone.

Despite the benefits of additional light while imaging provided by a ring light, it is best if this component can be avoided in the system. Ideally, the system is illuminated solely with coaxial, Köhler illumination to maintain uniform illumination and avoid imaging the light source directly. The Köhler source must be very bright since it is the sole light source and the system has loss due to rough/non-specular surfaces and multiple beamsplitters, but achieving such is possible.

### 4.3 Sensor Selection

Two specific sensors must be selected: one vision sensor, to collect information from the interferometer, and one sensor with ROI capabilities, to collect the final 2D image.

#### 4.3.1 Vision Sensor

Choosing the vision sensor requires three main considerations: the size of the sensor, the speed of operation, and the type of operation/processing that occurs. The optimal sensor has a large number of pixels, as close to the original 2000x2000 pixel requirement (Section 1) as possible. The speed should be quick due to the limited time for operation of both the imaging and interferometer scans (Table 1). Lastly, the type of processing done by the sensor should be simple and not user-intensive for ease of implementation.

There are three specific creators of vision sensors applicable to this system: Fraunhofer, Prophesee, and Inivation. They each produce vision sensors of varying sizes and with varying capabilities. Fraunhofer produces the sensor with the largest number of pixels, allowing for a field of view around  $\frac{1}{4}$  the area of the original system requirement. Other sensors are smaller, with Prophesee's ONBOARD the next largest, and Inivation's DAVIS 346 and DAVIS 240 sensors the smallest (Figure 45).

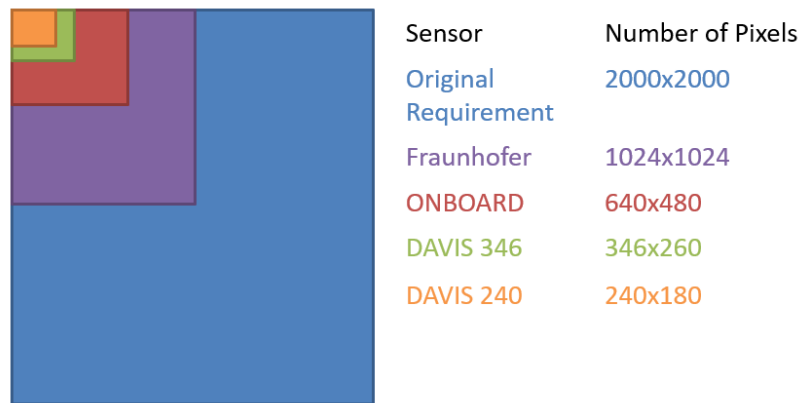


Figure 45. Comparison of the sensor size for the original requirement and various vision sensors.

The Fraunhofer sensor is advantageous because its field of view is so large. However, Fraunhofer does not offer a product that has specific functions. Its “Vision System-on-Chip” has certain inherent features, but requires “customer-specific further development of existing hardware and software components,” (Fraunhofer IIS) which indicates more input from the customer and likely a more difficult, expensive, and time consuming process. The vision sensor for this project must be a COTS component, and this sensor from Fraunhofer is not strictly that.

Instead, the sensors are narrowed down to Prophesee's ONBOARD and Inivation's DAVIS 346, since they are the next-largest sensors that are available off-the-shelf. Prophesee is preferred for its larger sensor (640x480 pixels compared to 346x260 pixels).

The Prophesee and Inivation sensors operate in similar ways. They both have a bias-type setting where the required change in irradiance to trigger an event can be set. They output events only when irradiance changes based on this bias and otherwise output no events.

Table 4. Comparison of speed and event output type for top two contenders: Prophesee ONBOARD and Inivation DAVIS 346.

Sensor	Speed (MEPS)	Binary?	True Irradiance?
Prophesee ONBOARD	25	Y	Y
Inivation DAVIS 346	12	Y	N

They differ, however, in their speed and the type of information they can output. ONBOARD operates at almost twice the speed of DAVIS (Table 4). Speeds for vision sensors are not quoted as “frames per second” but rather mega-events per second (MEPS), or how many million pixel

events can be recorded in each second. In this way, ONBOARD again is advantageous over DAVIS.

Additionally, ONBOARD can output binary events as well as true irradiance, while DAVIS 346 is only able to output binary events, where irradiance is interpolated between frames taken at 30 fps. This interpolation is not possible at the speeds the system will run – in 40 ms, only 1-2 frames total occur so accurate interpolation is not possible. Although true irradiance is not used in the current system design, it has potential to be used in a system variation.

Prophesee's ONBOARD vision sensor is selected for the system due to its larger field of view, faster speed, and ability to output true irradiance values.

#### *4.3.2 Imaging Sensor*

Like the vision sensor, the imaging sensor should operate as quickly as possible. Ideally, the sensor allows individual pixels to be accessed at random so when they come into focus they can immediately be read out. Such operation is theoretically possible with a CMOS detector architecture since pixels can be individually accessed, unlike CCDs in which the smallest feature accessed is a single row. Although it is theoretically possible, no current CMOS detectors have this function. CMOS architecture is trending toward this functionality, as block ROI and skip logic develop, which allows parallel operation of pixel groups as small as 32x32 pixels (Schrey). Even further, a recent development by CEA Leti was released which allows parallel operation and significant programming at the level of 16x16 pixels (Blackman).

Since random individual pixel access is not yet available, the sensor operation relies on collecting ROIs at various depths. Often times, ROIs are selectable by a user actively interfacing with a camera through software and a GUI on a computer. This is slow and user-intensive, so the sensor must have a dynamic, programmable ROI feature where a large number of ROIs can be programmed and triggered as the system scans.

Various sensors were considered, including as the Allied Vision “Bonito Pro,” the Optronis “CamRecord,” and the JAI “GO-2400C-PGE.” These three sensors represent the three distinct modes of operation available with ROIs. The Bonito Pro collects a single ROI. Switching between ROIs requires restarting the camera which takes a significant amount of time (Allied Vision). The CamRecord has greater flexibility depending on the software implemented. Up to 16 ROIs can be selected and the origin of the ROI can be changed while recording. However, the ROI shape and size cannot be changed, which is quite limiting if applied to the system (Optronis).

The GO-2400-PGE sensor is the only sensor found for this research that performs as required. It has a dynamic ROI ability, where it can switch between up to 128 different ROIs. These are programmable and can be triggered to record at required times (JAI). Because this sensor is the single COTS component that works as required, it is chosen for the system.

## 4.4 Final System Performance

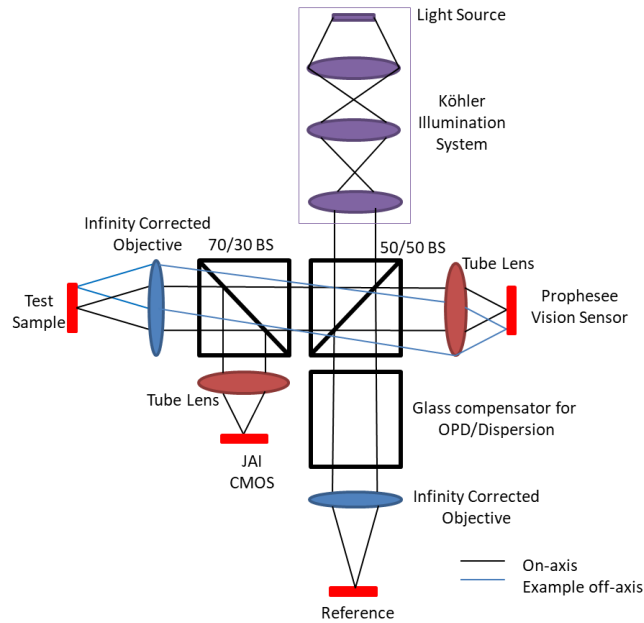


Figure 46. Full layout of final system, detailing all components necessary.

The final system is an interferometric microscope which has been modified by the insertion of a second beamsplitter to create an imaging arm. Köhler illumination is used to illuminate the interferometer and imaging arms of the system, which is fully conceptualized in Figure 46. The main components of the system – including objectives, tube lenses, sensors, and beamsplitters – are detailed in Table 5.

Table 5. Components used in the final system.

Component	Name	Provider/Part Number	Description
Objective	10X EO M Plan Apo Long WD Infinity Corrected	Edmund Optics, 59-877	10X, Infinity Corrected
Tube Lens	MT-1 Accessory Tube Lens	Edmund Optics, 54-774	Matched to objective
BS 1	50/50 Cube Beamsplitter	Edmund Optics, 54-823	50/50, Cube, C-Mount
BS 2	30R/70T Plate BS	Edmund Optics, 49-684	70/30, 3mm thick plate, BK7 glass, C-Mount
Compensator Plate	35 mm sq x 3 mm thick uncoated Bk7 window	Edmund Optics, 49-470	3 mm thick plate, BK7 glass
Vision Sensor	ONBOARD	Prophesee	640x480 pixels, 15 $\mu\text{m}$ pixels
Imaging Sensor	GO-2400C-PGE	JAI	1936x1216 pixels, 5.86 $\mu\text{m}$ pixels, up to 128 programmed ROIs

The system meets geometrical requirements. The 0.28 NA microscope objective allows pixel resolution requirements to be met. In practice, pixel resolution may become limited due to vibration from the scanning mechanism. A low vibration scanner is chosen to limit this effect, and efforts can be made to implement vibration isolation if necessary. Also, the final pixel resolution is made slightly worse by using a 10X rather than 12X microscope.

The system has no inherent limit on scan length, so the depth of focus (500  $\mu\text{m}$ ) is easily attainable. Depth resolution is easily met by an interferometer and is adhered to by grouping ROIs based on their “best focus,” where the measurement of a pixel occurs no more than 3  $\mu\text{m}$  from its true best focus location. Because the ONBOARD sensor is used, with 640x480 pixels with 15  $\mu\text{m}$  pixel pitch, the object field of view is 0.6 x 0.8 mm.

Collecting a true color image with this system is much simpler with the imaging arm separate from the interferometer. Interference does not impact the final image, and the discussion about careful selection of a source based on its spectrum (Section 2.4) is no longer relevant; the imaging arm operates without added difficulties from an interference pattern. The source spectrum does not need to be carefully chosen. It simply must be “white light” so a true color image can be produced. The separate imaging arm also avoids issues of sudden brightness changes that occur with interference fringes which had the potential to affect the brightness and color of a final image. Irradiance does not change quickly in the imaging arm, but appears as any typical image does on a sensor: certain parts will be in focus with other parts defocused, and the image will change while scanning so different parts come in to focus. Removing the interference pattern from the imaging arm makes collection of the final image much simpler, such that imaging operates in a more typical and simpler manner.

#### *4.4.1 Speed of Operation*

The system operates slower than requirements demand as two sensors and two scans in sequence are used. The speed of the system is determined by the speed of the vision sensor, the speed of the imaging sensor, and the speed of processing to create ROIs and compile the final image.

The Prophesee ONBOARD vision sensor’s speed of operation is determined by the number of pixels in the detector, the estimated number of times each pixel triggers an event, and the speed of the vision sensor. The detector has 640x480 pixels, or at total of 307,200 pixels. The bias is set to output events with certain changes in irradiance, and for a typical pixel it is desired that 3-5 events are outputted in the fringe region. If a pixel has particularly high reflectivity and therefore higher fringe contrast, more events may be outputted. For a low reflectivity pixel, fewer will be outputted, but the bias will be determined such that low reflectivity pixels output 1 event to find best focus. With an average 3-5 events per pixel, it is expected there will be about 900,000-1,500,000 events for the entire pixel array over the entire scan depth. The vision sensor operates at a maximum 25 MEPS and the pixel array is mapped in 35-60 ms.

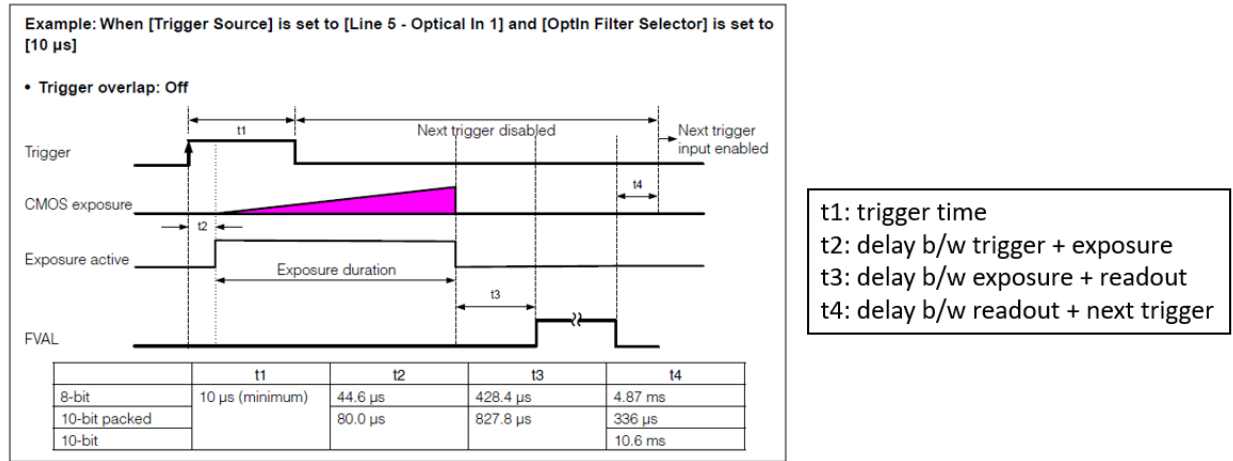


Figure 47. Breakdown of speed of JAI GO-2400C-PGE operation with trigger for ROIs (JAI).

The speed of the JAI GO-2400C-PGE imaging sensor is based on the triggering operation switches between ROIs. The time for a single ROI – from triggering the ROI to the delay before the next triggered ROI – comes from a combination of the exposure time, readout time, and four different delays (Figure 47). The minimum exposure time is 38  $\mu$ s and the delay between the trigger and exposure overpowers the trigger time minimum (t1). The readout time comes from the size of the ROI since readout is based on line frequency (time to read out a single row of the sensor), which is 12.445 kHz for 8-bit and 24.2425 kHz for 10-bit. So, the total time for one ROI becomes

$$ROI \text{ speed: } t2 + \text{exposure time} + t3 + (\text{line frequency})(\# \text{ of lines}) + t4. \quad (14)$$

The shortest readout for a single ROI occurs when the ROI is a single line and minimum exposure is used. This readout takes 1-12 ms (based on if the data is 8-bit, 10-bit packed, or 10-bit). The maximum readout for a single ROI occurs if all lines are readout and the ROI contains the full pixel array. Although this ROI is not practically used, it represents the upper extreme for ROIs and requires 21-44 ms. The minimum number of ROIs to read out is 1 and the maximum is 167 (based on the scan depth and optical resolution).

There is an unknown number of ROIs that are created and each has an unknown size, which causes the speed of the imaging sensor operation to change. Based on the speed for a single line ROI, it is clear the imaging sensor will not operate as quickly as required. Despite this, the JAI-GO-2400C-PGE is still used because of its promising type of operation. Random pixel access is a focus of research currently and is expected to develop greatly in the coming years such that in the near future the time requirements can be met.

It is expected the processing during the operation should not take significant time compared to the sensor operation times. The processing of the depth map from the vision sensor to create ROIs to input into the imaging sensor is the most involved. This process may be time-consuming to automate, but it is expected an algorithm can be created which optimizes the size and locations of ROI to record in the shortest amount of time. The algorithm itself is expected to operate relatively quickly due to simple operations. There is also very little delay between event occurrence and output, so there is potential the algorithm can operate to group ROI while the scanning occurs.

The final processing to create a single in-focus image is quite simple. When each ROI is outputted, the relevant in-focus pixels are saved and copied into their proper locations in a separate array. This is done for each ROI to fill the entire array with in-focus pixels.

Overall, the speed of operation does not meet requirements. The vision sensor takes around 40 ms and the imaging sensor will take far longer, meaning the speed will be much slower than the requirements allow. However, this approach is still pursued because of its promising future use and increased speed with near future technology. The system created with these parts works as a prototype and although it sacrifices speed, it maintains other requirements and works as a proof-of-concept for this new type of system.

## **5. Modeling**

Portions of the system were modeled in FRED to better understand performance. There are multiple nuanced factors to address to successfully model white light interference phenomena in the software and produce a cohesive, full understanding of the concerns of modeling WLI systems. Such analysis clarifies what features in the model result from interference effects or real phenomena of the system and which are artifacts of the modeling. The illumination and imaging portions of the system are also modeled to ensure they operate as expected.

### **5.1 Coherence in FRED**

To model interference in FRED, a coherent source is used. The incoherent white light source is modeled such that its wide spectrum is discretized into many individual wavelengths, each analyzed separately to create fringe patterns. The fringe pattern irradiances are summed to represent the fringes of the white light source. This method of representing white light interference is the same method used in describing temporal coherence effects (Section 2.1). Each individual wavelength is perfectly coherent with itself and is raytraced in FRED independently of each other wavelength. The results of many coherent raytraces for single wavelengths added together produce results which in theory mimic the interference of a broadband source.

Modeling coherence in FRED differs greatly from incoherent modeling. With incoherent modeling, each ray propagates in its specified direction with a specified power, only changing direction or other properties when it contacts a surface, such as a lens. When a ray contacts an analysis surface, its power is measured at the exact location it intersects, in the single pixel of the analysis surface with which it comes in contact.



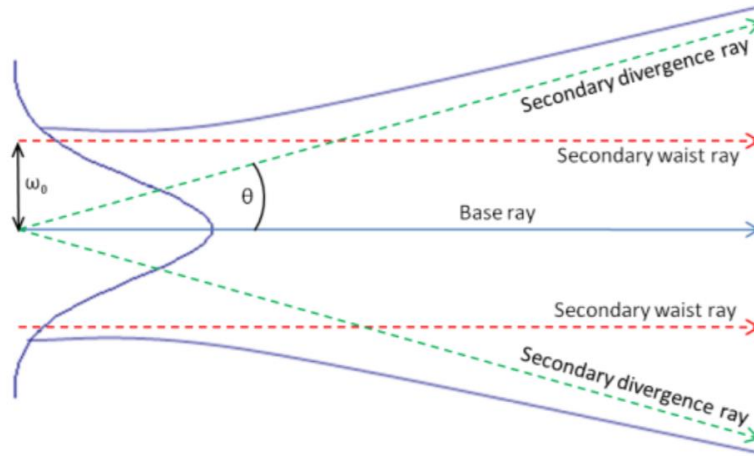


Figure 48. Gaussian beam and 5 of the 9 rays defining it. The other four rays are divergence and waist rays into and out of the page (FRED Application Note: Modeling Coherence).

When modeling coherent light, each “ray” created by the user produces a small Gaussian beam (Gausslet) which is represented by 9 rays total: one central “base” ray, 4 waist rays, and 4 divergence rays. The Gausslets add complications to raytracing. For a Gausslet to continue propagating, the center ray must hit the surface, and the other 8 rays must hit the surface or the “mathematical extension of the surface” (FRED Application Note: Modeling Coherence) if the physical edge is continued. There are cases, such as at the top of a sphere, when all rays will not intersect an extended surface and in these cases the Gausslet is not propagated.

Irradiance values measured at the analysis surface are very different than when modeling incoherent light. Gausslets have spatial extent, as opposed to incoherent rays which act only on a single point. This causes the power from each Gausslet to be spread over an infinite area based on the Gaussian distribution. So, rather than one ray going to one pixel, as in incoherent modeling, each Gausslet contributes to the irradiance at every pixel based on its distribution.

## 5.2 Noise in Model: Gausslets

When the WLI system is modeled in FRED, the Gausslets become the source of many types of noise that appear. Gausslets create imprecision in modeling with a Bayer filter and cause fluctuations in irradiance based on model characteristics, such as Gausslet overlap factor, source ray density, and propagation distance.

### 5.2.1 Bayer Filters

A Bayer filter is present in a traditional color image sensor and therefore is modeled to represent the imaging sensor of the system. A Bayer filter is an array of red, green, and blue filters which align with individual pixels (Figure 49). Color filters allow for color imaging when simple image processing is performed on the raw image.

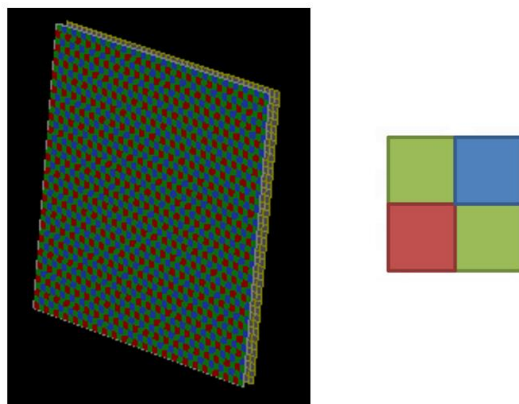


Figure 49. Bayer filter built in FRED (a) and a representation of a single Bayer filter "pixel" (b) which corresponds to four detector pixels.

Modeling a Bayer filter in FRED is problematic. When a Gausslet contacts the Bayer filter, spectral characteristics of the Gausslet change based on the intersection of the base ray. If the base ray goes through a red pixel in the Bayer filter, the entire Gausslet is filtered to the red spectrum. The Gausslet extends to all pixels, contributing irradiance in the red spectrum to each. In a real system, the spectrum accepted by each pixel is the spectrum of its filter, so this spread of color from individual Gausslets is an unrealistic artifact of the model. This modeling effect occurs for all Gausslets, making it a potentially significant artifact of the modeling. The degree of noise induced by this effect is dependent on the characteristics of the Gausslets.

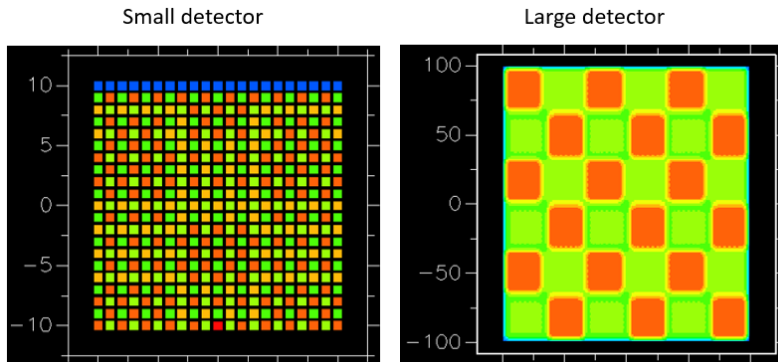
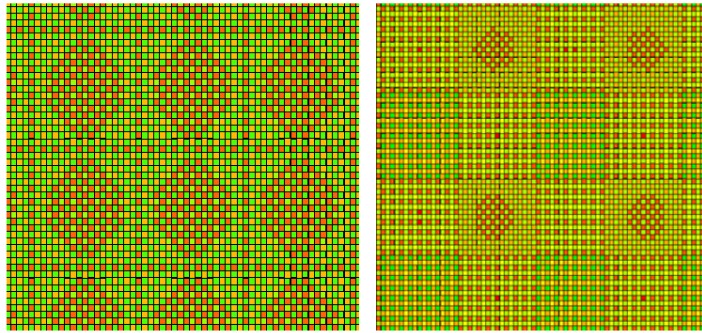


Figure 50. Irradiance on detector with Bayer filter for small (a) and large (b) detector. A small detector is modeled accurately, but the large detector is affected by aliasing.

In practice, using a Bayer filter with a coherent model requires trial-and-error to determine the proper source properties to avoid significant artifacts. Checking for these artifacts is done in the model with a perfectly flat, untilted object at its best focus, so constructive interference occurs across the object for all wavelengths. Such a setup produces different irradiance for the different color pixels (red, green, and blue), but consistent irradiance between pixels of the same color. When a small detector and corresponding Bayer filter array are used in the model, irradiance maps generally show pixel-level behavior and an accurate model of the Bayer filter. When the detector and Bayer filter increase in size, the effect of color spread from the Gausslets begins to dominate and aliasing occurs. Changes in irradiance do not occur at the pixel level but on a larger scale,

creating a pattern with features at around 20x20 pixels which indicate aliasing due to undersampling (Figure 50).



*Figure 51. Two example patterns showing effects of coherent modeling with a Bayer filter. Increasing ray density no longer improves upon these patterns.*

Aliasing effects are mitigated by changing ray density of the source, such that there are more Gausslets. This allows for better sampling of the Bayer filter (Davies). The aliasing pattern is larger with fewer rays and gets smaller as ray density increases. There is a point where changing ray density no longer improves the effects of Gausslet interaction with a Bayer filter, and there is still some remaining noise (Figure 51).

Such artifacts do not noticeably occur in small sensors but cannot be completely removed when an analysis surface size is increased to a certain point. As such, modeling is altered to remove the Bayer filter. The alternative model requires three traces to be performed, one for each color of the Bayer filter. Adding results from the three raytraces together produces a result closely matching what is expected in a real world system. However, such modeling does not account for the spatial layout of a Bayer filter and may create a model with higher spatial resolution than can be achieved with a Bayer filter. However, such a difference between the model and real system can be accounted for and is better understood than the aliasing effects of the Bayer filter, making it a reasonable alternative in modeling.

### *5.2.2 Amplitude Fluctuations*

Gausslets create artifacts in the model in the form of irradiance fluctuations. A source with a “flat top” beam produces uniform irradiance across the object. The flat top is easily modeled with an incoherent source, but becomes warped by amplitude fluctuations in coherent methods (Figure 52). Each Gausslet has varying amplitude across the object and the sum of the Gausslets creates a ripple pattern which adds noise to the model. The fluctuations are small compared to the average amplitude and therefore does not have a large effect.

There may be small signals, however, when a larger amount of light is returned from the reference mirror than the test arm, such as when a low reflectivity, highly scattering object is imaged. For these cases, it is important noise is minimized in the model so the effects in the image are dominated by object features rather than noise.

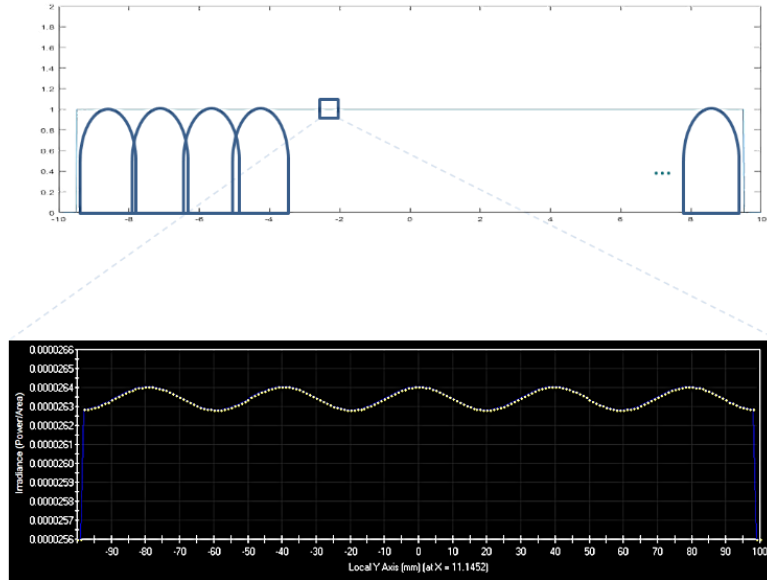


Figure 52. Ripples in model of a "flat top" beam which are caused by the combination of many Gausslets.

This ripple effect is altered by a few different parameters, the first of which is the “overlap factor.” The overlap factor is the “fractional overlap of the waists of adjacent Gaussian beams when created in a grid” (Photon Engineering). If the overlap factor is 1, the beam waist rays of one Gausslet are perfectly coincident with the beam waist rays of neighboring Gausslets.

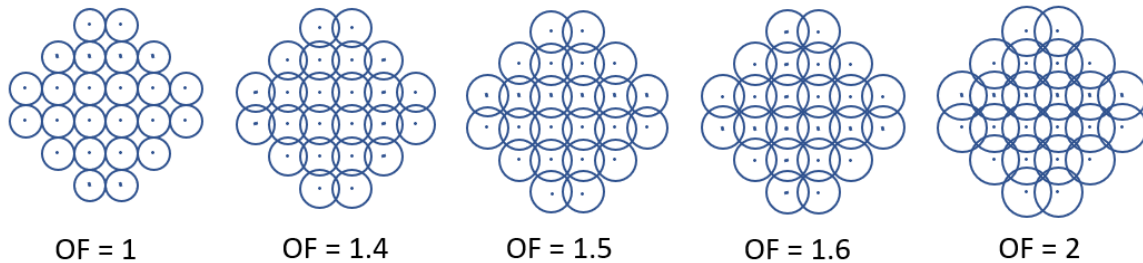
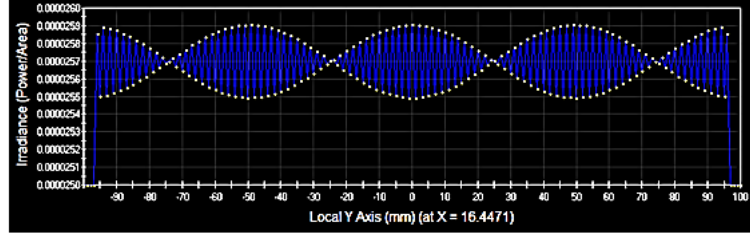


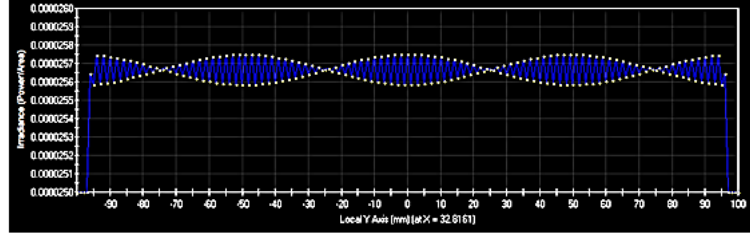
Figure 53. Visualization of overlap factor, showing base ray and Gausslet beam waists.

FRED suggests the use of an overlap factor of 1.4 to 1.6, visualized in Figure 53. When overlap factor increases, fluctuation in irradiance decreases. However, the “roll off” on the edges becomes shallower, creating a smaller area with uniform irradiance (Figure 54). The overlap factor is chosen to balance between these effects such that fluctuations are minimized to keep signal-to-noise ratio (SNR) high but also so the edges of the field of view are not dimly illuminated due to the roll off effect. An overlap factor between 1.5 and 1.6 is best for this model, which keeps the SNR between 250 and 500.

Overlap factor = 1.4



1.5



1.6

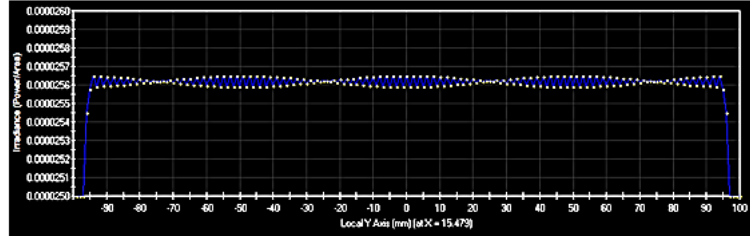


Figure 54. The effect of overlap factor in the magnitude of the fluctuations from individual Gausslets.

Overlap factor sets the beam waist location, so it must also be carefully chosen based on Gaussian beam characteristics. Gausslets perform best if they remain paraxial, based on the Gaussian equation for divergence half-angle  $\theta$  and waist radius  $\omega_0$ ,

$$\tan \theta = \frac{\lambda}{\pi n \omega_0}. \quad (15)$$

If Gausslets are non-paraxial, they are unable to “accurately sample optical components” as waist and divergence rays fail to “remain well-correlated with their base ray.” The paraxial approximation requires that  $\tan \theta$  is essentially  $\theta$ , and therefore  $\theta$  should be around 6 degrees or less. To ensure this is met, the waist must be “greater than or equal to about  $3\lambda$ ” and typically in practice is  $5-10\lambda$  (FRED Application Note: Modeling Coherence). Choosing overlap factor must account for this condition as well as the issues of fluctuations in irradiance.

Fluctuations in irradiance vary with ray density. This is demonstrated by keeping the light source the same size while performing traces with varying number of rays to represent the light source. Ray density is described by ray semi-aperture and is varied from 101 to 2001 rays as an example. As ray density increases, fluctuations decrease and SNR increases. However, increasing ray

density causes more significant “ringing” at the edges of the field, where the irradiance increases and decreases rapidly, like ringing from diffraction (Figure 55).

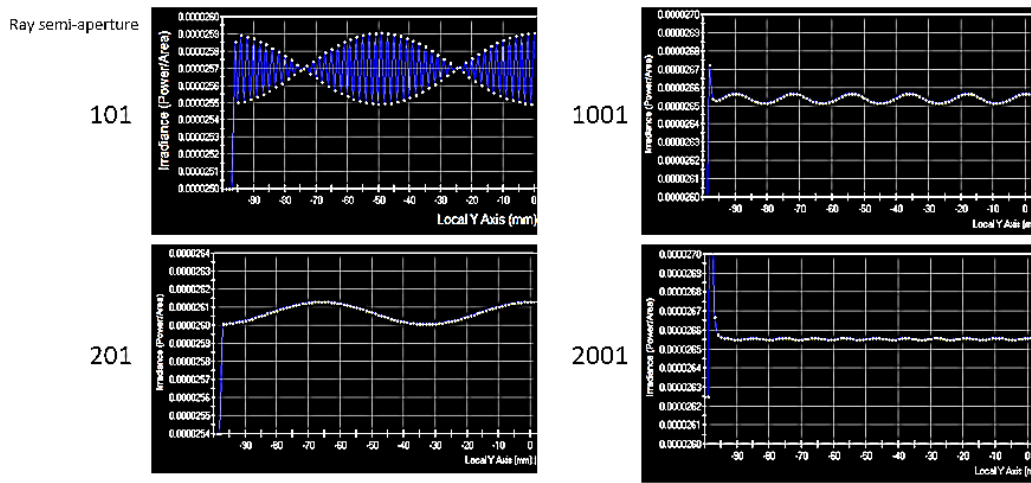


Figure 55. Fluctuations in irradiance over half the FOV for varying ray densities (ray semi-aperture of 51 to 2001). Note that as fluctuation amplitude decreases, ringing increases on the edges of the aperture.

A dampening factor can be introduced to the edges of the field of view to reduce the ringing effect. If this option is not selected, a mid-level ray density balances between fluctuations and ringing at the edges. This is generally trial-and-error based rather than driven by any specific formulas or rules, other than keeping the Gausslets paraxial by keeping grid spacing larger than  $2\text{ }\mu\text{m}$  and for best operation larger than  $5\text{--}10\text{ }\mu\text{m}$ , given visible light is being modeled (FRED Application Note: Modeling Coherence).

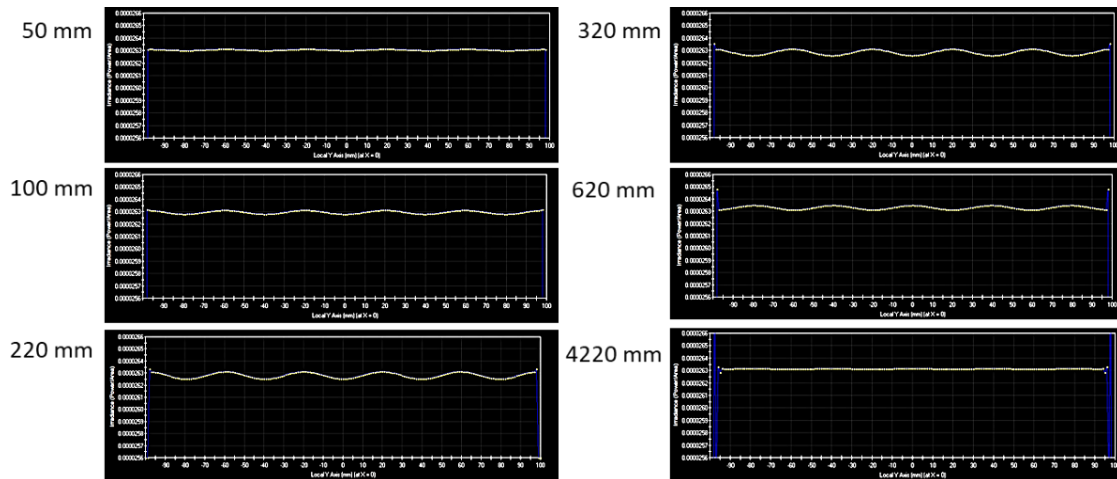


Figure 56. Irradiance plots showing the ripple effect and how it changes with propagation distance.

Amplitude of fluctuations also changes with propagation distance. Initially, when propagation distance increases, fluctuations increase. There is a point where fluctuations decrease again and ringing appears on the edges of the field (Figure 56). This likely indicates that there are optimal raytrace distances and after such distances, the field should be resampled, where “a new set of coherent rays [reproduce] the current scalar field” such that the “new rays sum up to generate the same field, but the freshly synthesized beamlets have redefined waists and divergence angles”



(FRED Application Note: Modeling Coherence). This creates a set of “well-behaved” Gausslets, replacing the old Gausslets which may have become less well-behaved after propagating some distance through the system. In general, resampling is used when any of the effects described are observed and become significant through the model. Often, resampling is a simple and effective solution to minimizing fluctuations in irradiance, especially when paired with well-behaved Gausslets.

### 5.3 Artifacts from Discrete Wavelength Modeling

There are artifacts in the FRED model from modeling the system with discrete wavelengths rather than a continuous white light spectrum. Discrete wavelengths cause fringes to have a “resurgence” well beyond the zero OPD region. The central fringe pattern around zero OPD is pronounced and the signal goes down to the incoherent level beyond the zero OPD region; however, at higher OPDs, the fringe pattern revives due to the false periodicity that arises from the selected discrete wavelengths. Ultimately, the resurgence location depends on the number of wavelengths, the spacing between them, and the distribution of these wavelengths.

The resurgence in modeling differs greatly from the real effects that are observed with white light interferometry. The real interference pattern comes from a continuous spectrum and resurgence does not come about since a wide bandwidth spectrum is used. Discrete wavelengths break the characteristic wide spectrum and instead represent a wide spectrum with many discrete spectra added together.

To understand the effects of this in the model, a flat, tilted object is used. It is tilted such that the difference in height from one end of the object to the other is the full scan depth, in this case 500  $\mu\text{m}$  (Figure 57). With a continuous spectrum white light source, fringes appear in the center of the object, where OPD is zero, and the rest of the object is at its incoherent level.

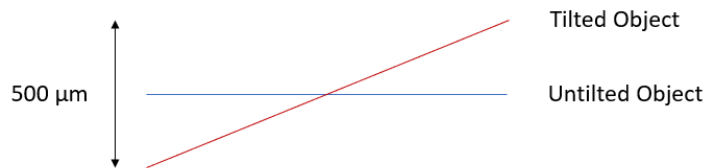


Figure 57. Schematic of a tilted object, tilted to capture all possible heights of an object in a 500  $\mu\text{m}$  depth scan.

When modeled in FRED with a discrete number of wavelengths, this is not achieved. Instead, at a certain point in the object, fringes have a resurgence and fluctuations in irradiance occur along the edges of the object (Figure 58). The point where resurgence occurs must be “pushed out” far enough so resurgence will not occur for the 500- $\mu\text{m}$  depth required for the system *DOF*. At this point, the discrete model accurately represents the continuous function.

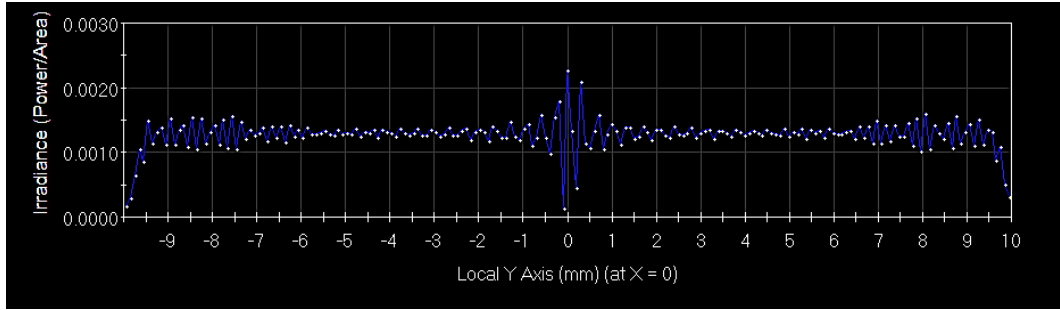


Figure 58. Artifact in the form of resurgence in fringe structure in FRED due to use of discrete wavelengths.

The resurgence of fringes causes fringes to reappear on the edge of a modeled tilted surface, and also to appear on features that are out of focus, if they are at the proper depth where resurgence occurs (Figure 59). This indicates in the model that such objects are at their best focus locations and signals events on the vision sensor, which would not occur in a real system and shows the model is not accurate.

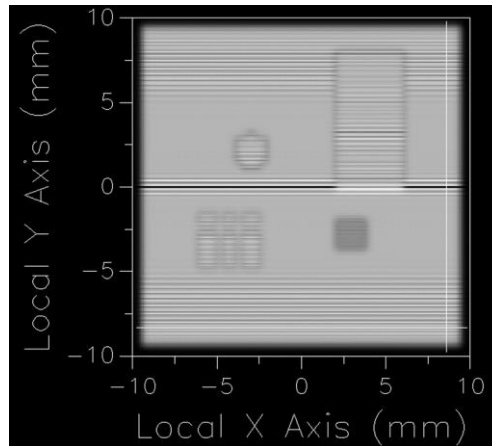


Figure 59. Fringe resurgence as seen on out-of-focus features of the object and on the top and bottom edges of the field.

Improvement is achieved by increasing the number of wavelengths, but each additional wavelength traced adds significant time to the model computation. The best solution uses the fewest number of wavelengths to remove resurgence from the required depth.

The FRED default spacing between wavelengths is linear, so all wavelengths are evenly spaced. The number of discrete wavelengths can be increased to push out the location for fringe resurgence. To achieve a height range  $X \mu\text{m}$  without fringe resurgence (equivalent to  $2X\text{-}\mu\text{m}$  OPD range in a double pass interferometer), approximately  $4X$  discrete wavelengths must be traced in the visible spectrum ( $0.4 - 0.7 \mu\text{m}$ ), which was empirically derived during this modeling process. For a depth of  $500 \mu\text{m}$ , approximately 2,000 wavelengths must be traced. For the current model, tracing 51 wavelengths can take 30-60 minutes, and speed decreases linearly with number of wavelengths. It is worthwhile to seek a modeling method using fewer wavelengths.

Allowing some random offset to perturb each wavelength traced is explored. Each wavelength has a nominal value, and a random offset within a certain range around this nominal value is allowed so the wavelengths are not evenly spaced. Resurgence occurs when fringes from different wavelengths begin to align again, and it is expected that if the wavelengths are not evenly spaced,



the resurgence will occur after a larger OPD than if they are evenly spaced. In practice, however, this is not the case, and allowing a random offset caused resurgence to happen sooner and introduced more unpredicted fluctuations in irradiance into the system (Figure 60). Ultimately, this adverse result is to be expected, such that rather than random offsets, one must appropriately choose the offsets between the discrete wavelengths.

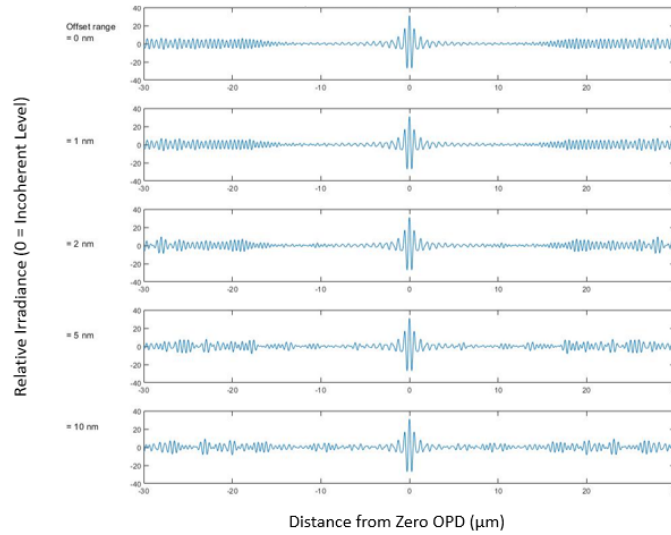


Figure 60. Fringes for 31 discrete wavelengths, with various ranges for random offsets of wavelengths.

Thus, the spacing of wavelengths is further explored by allowing for logarithmic rather than linear spacing. This choice somewhat changes the shape of the fringes, but not significantly, and reduces the number of wavelengths that must be traced. With logarithmically spacing the resurgence is pushed out and there is a 25% improvement over linear spacing (Figure 61).

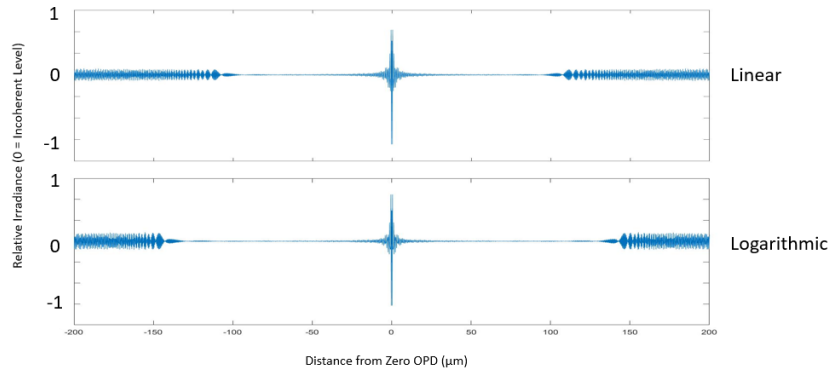


Figure 61. Fringes from (a) linearly and (b) logarithmically spaced wavelengths.

The logarithmic spacing improvement make sense when the wavelengths are considered as red, green, and blue components, as with a Bayer filter. If each spectral range contains an equal number of wavelengths with equal spacing, longer wavelengths produce resurgence that is pushed out further than when shorter wavelengths are used (Figure 62). Logarithmic spacing takes advantage of this and spaces shorter wavelengths closer together, pushing resurgence outward for these wavelengths. This allows for fewer wavelengths to be traced overall.

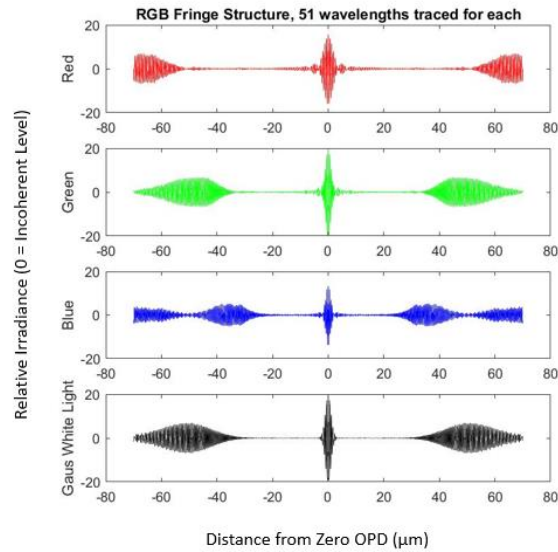


Figure 62. Fringe structure and how it varies with wavelength range.

The amplitude function of the source also affects the location of fringe resurgence. This is explored in general by comparing uniform amplitude to Gaussian amplitude, where the peak of the Gaussian function occurs at 550 nm and is wide enough to encompass the visible spectrum. The fringe resurgence is pushed out further for a Gaussian amplitude function in comparison to uniform function (Figure 63). The amplitude relates to a source spectral density, and this observation indicates that when a real spectrum is used in the model, it will perform better than the uniform amplitude model, since it will more closely resemble a Gaussian function. Specifics are dependent on what source is used, but indicates that even fewer wavelengths are needed to properly model the fringe structure.

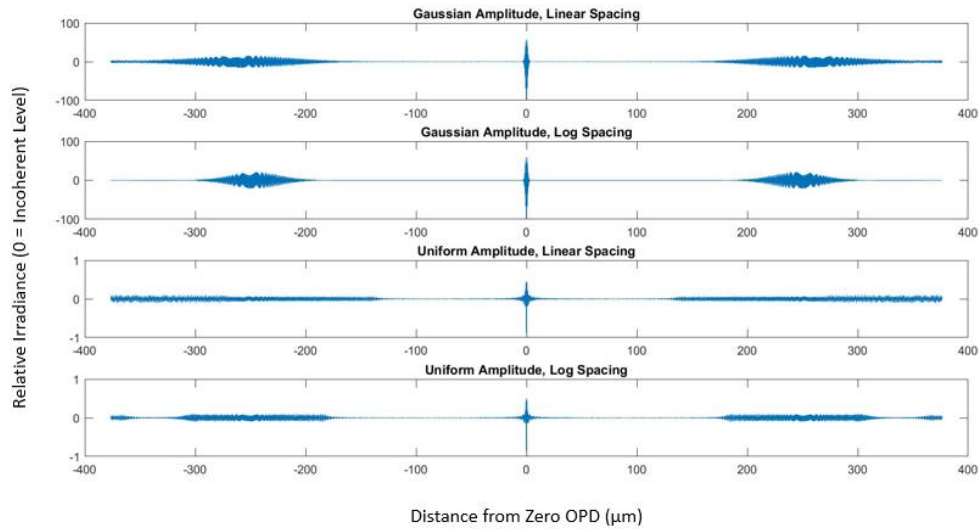


Figure 63. Fringe resurgence for combinations of Gaussian amplitude, uniform amplitude, linear spacing, and logarithmic spacing.

## 5.4 Roughness

The type of surfaces measured with the device must be considered for a complete model of the system in FRED. Fringes shown up to this point are for perfectly specular surfaces and as such the fringes appear very smooth and uniform across the objects. However, the system will measure general objects which can be optically rough. When modeling incoherently, surfaces are made to scatter based on the specific bi-directional scatter function (BDSF) of the object. However, scattering surfaces cannot be used with coherent modeling in the software. There are no models that allow for this at this time, so the software removes all phase information of the rays so scattering cannot be used with coherent modeling.

Instead, surface roughness is used, which “[introduces] an element of randomness into the ray intersection/interaction process during the raytrace” (Photon Engineering). The surface normal vector is perturbed based on user-defined roughness specifications. The normal is perturbed more often upon ray intersections on rougher surfaces. Also, the normal is perturbed to a larger degree on a rougher surface.

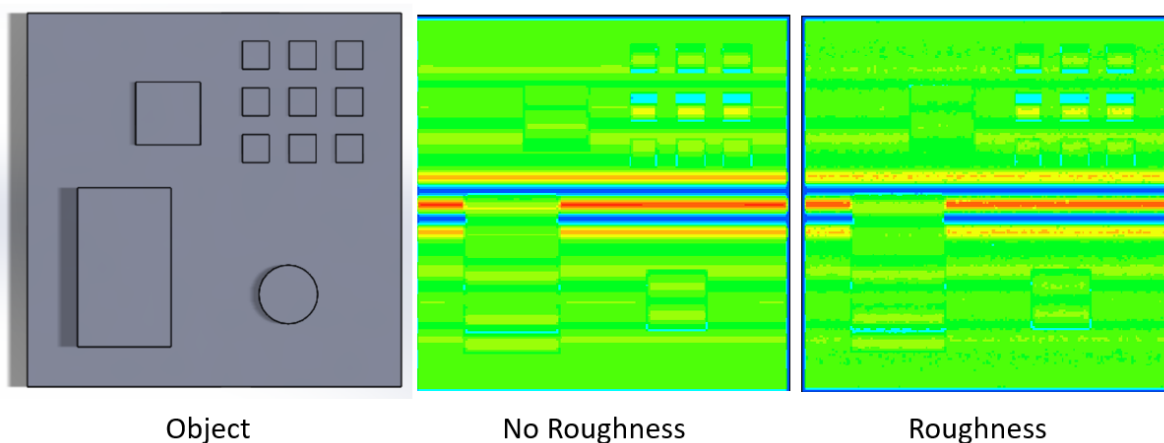


Figure 64. Modeling optically (b) specular and (c) rough surfaces in FRED as the system must work for a (a) general object.

In the interferometer, roughness changes the appearance of fringes, so they are not as even and linear on the object surface (Figure 64). These fringes match what a generally rough object produces in real life (Figure 65). It is especially important to use roughness modeling to determine the amount of light lost by a rough surface to know how much light is expected in the test arm of the interferometer and in the imaging path.

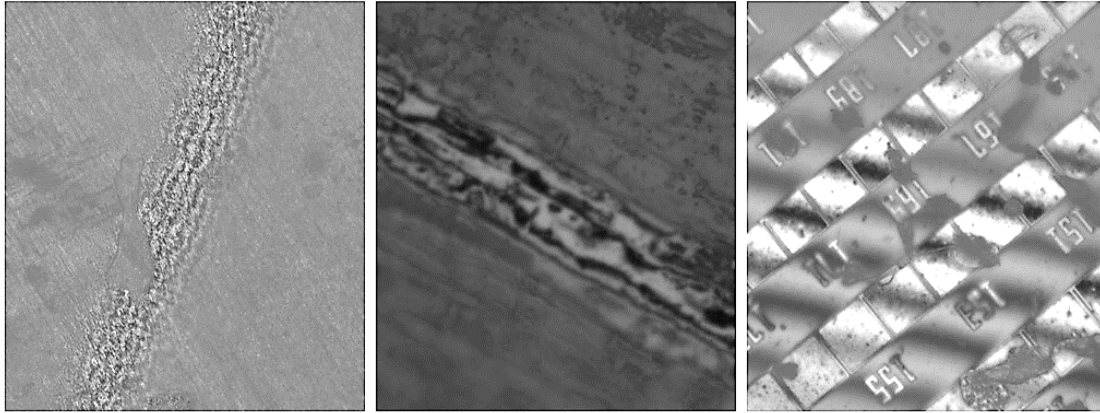


Figure 65. Fringes from various semiconductor chip surfaces, measured with a Zygo WLI.

Addressing these considerations creates a model that accurately represents fringe patterns that are expected in a laboratory setting. However, doing so is computationally intensive, particularly due to the number of wavelengths required to trace. In modeling this system, fewer wavelengths are traced and analysis is done over a narrower depth range. The model confirms fringe formation at zero OPD and the lack of fringes at higher OPD, which is the main purpose of the interference model.

## 5.5 Illumination Model

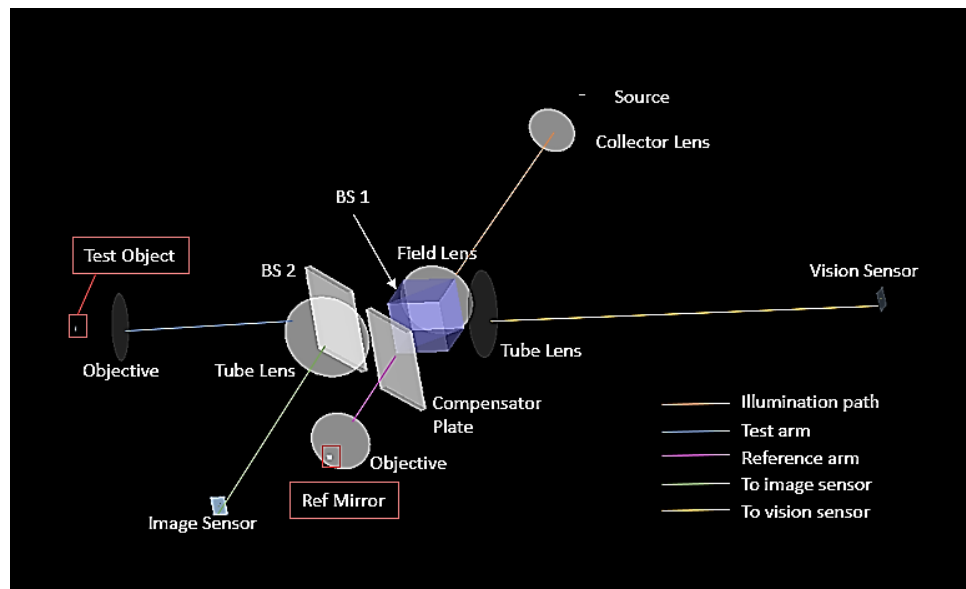


Figure 66. Layout of system modeled in FRED.

The system models three major components: the illumination conditions, the imaging conditions, and interference patterns (Figure 66). The process of modeling to produce accurate fringes has been exhaustively detailed, but the modeling is incomplete without considering illumination and imaging. The illumination system and imaging arm are both modeled to confirm their operation.

The illumination model allows the Köhler illumination scheme to be tested. Köhler illumination is used to produce uniform illumination of samples by interchanging the spatial and angular source emission distributions. As previously described, three lenses are used to create a Köhler illumination scheme: a collector lens, a field lens, and a condenser lens. The condenser lens is simply the microscope objective, which has a focal length of 20 mm. The field lens focal length is chosen based on physical space in the system and is set to 150 mm to be long enough to work with the system but still short to keep the system compact. A 150-mm focal length lens specifically is chosen because it is available as a COTS component. The collector lens focal length is chosen to allow the light source to be imaged to the back focal plane of the compensator so the size of the image matches the aperture size. So, the collector lens focal length is set to 35 mm.

The spacing of the lenses is known (Figure 40) and is modeled in FRED to confirm proper illumination. The illumination system is created using the characteristics of the light source, which has a diameter of 6.35 mm and a half-angle spread of 10.8 degrees (Advanced Illumination). Once included in the FRED model, a raytrace is done. Analysis is performed at the reference surface and object, where no interference occurs. So, an incoherent raytrace can be done. When properly set up, both the test object and reference mirror have uniform lighting (Figure 67). The difference in the irradiance levels is due to the chosen reflectances and transmittances of the beamsplitters in the system.

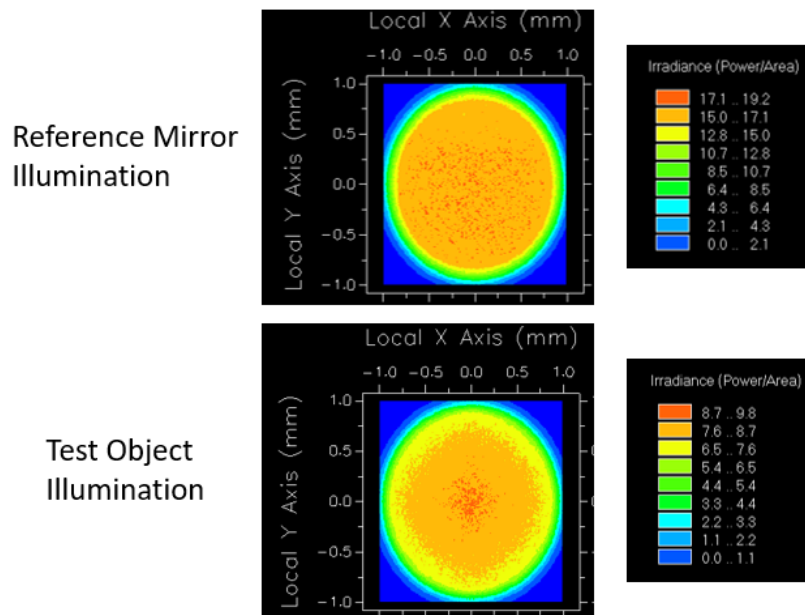


Figure 67. FRED modeling results for illumination on reference mirror (a) and test object (b).

The analysis surface in FRED is larger than the field-of-view of the system in order to find where the irradiance begins to fall off and where non-uniform color appears. Both occur at the far edge of the spot created by the illumination system, which is far outside the range of the system field of view. An object centered within the spot is uniformly illuminated in both arms. The illumination plots appear slightly differently. As expected, the irradiance on the test object is half that of the reference mirror when a 50/50 beamsplitter is placed as the second beamsplitter in the test arm. The irradiance increases when modeled as the system will operate, with a 70/30 beamsplitter instead. When the illumination on each surface is normalized and the two are subtracted, it is clear

the illumination conditions are nearly identical in the model (Figure 68). There are slight differences but are around 2% of the uniform irradiance on the test object and are considered negligible.

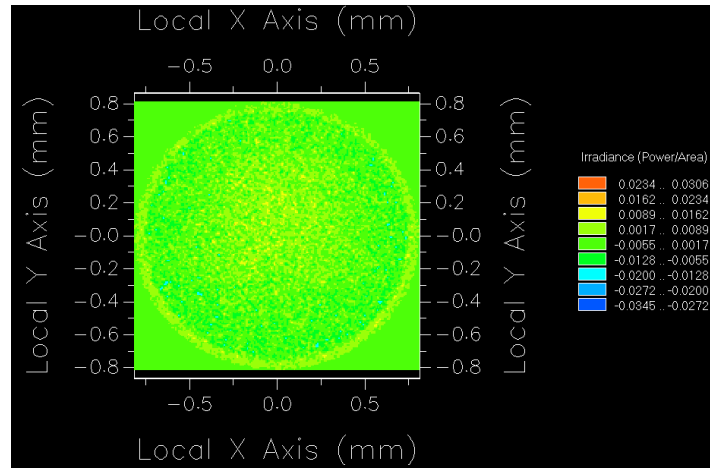


Figure 68. Relative difference in illumination between the reference mirror and test object.

The model verifies the design of the Köhler illumination system will work and provides further quantitative detail on how uniform the lighting becomes with this method.

## 5.6 Imaging Arm

The imaging arm operation is simple to confirm in FRED. Because the imaging arm does not use interference phenomena, it can be traced with incoherent light, just as the Köhler illumination setup. The FRED model confirms operation of the imaging arm, and that objects at their best focus location are sharp, while features outside of the *DOF* range inherent to the optical system are blurred. In Figure 69, this is shown clearly as the 3x3 square features are in focus with sharp, crisp edges and uniform irradiance, while other features of the object have blurred edges. Also note the background of the object was purposefully set to not reflect any light so these features can be seen more clearly.

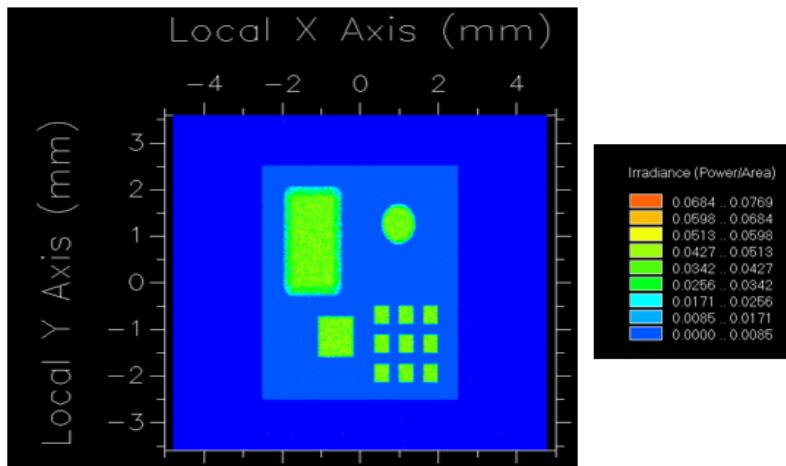


Figure 69. Sample object imaged with FRED model. 3x3 square pattern at best focus location.

The object is modified to include narrow rectangles, which are representative of wires that may populate the chip. Wires of 7.5-, 5-, and 3- $\mu\text{m}$  thickness are included, and all three are readily imaged (Figure 70). The 3- $\mu\text{m}$  wire image must be clear since 1.5- $\mu\text{m}$  object pixels have a Nyquist frequency of 3  $\mu\text{m}$ . Imaging this wire shows the system is operating as required. However, it is noted that objective and tube lens specifics are not known (i.e., the information is proprietary) and therefore exact performance cannot be included in the model. In this way, it is possible that the performance may be degraded by a small amount in the real system.

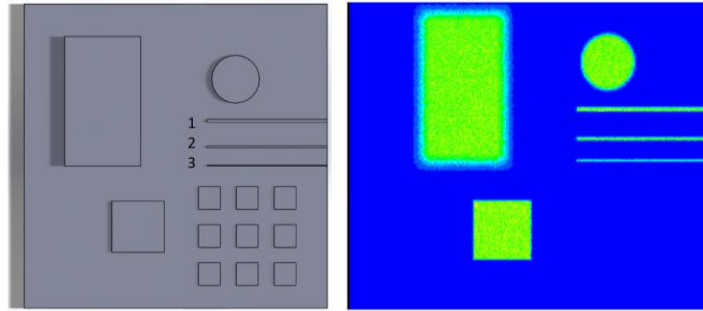


Figure 70. An object with three "wires" of thickness (1) 7.5  $\mu\text{m}$ , (2) 5  $\mu\text{m}$ , and (3) 2.5  $\mu\text{m}$ . All are imaged in the FRED model when at best focus.

Considering the factors relating to interference phenomena, illumination conditions, and imaging conditions allows accurate modeling of the system. Specifically, the process of modeling interference and controlling many factors (raytrace setup, effects of Gausslets, artifacts of use of discrete wavelengths, and roughness) allows for the creation of an accurate interference model which shows the promise of fabricating the system. A comprehensive study into the challenges in modeling white light interferometry in FRED is not only helpful with this project but is useful in general when working with coherence in FRED. Additionally, modeling the imaging and illumination sub-systems confirm overall operation of the proposed system.

## 6. Conclusions

A WLI system is conceptualized, designed, and modeled in preparation for creating a prototype which will create 2D in-focus images of 3D objects. The WLI system design meets nearly all defined requirement (Table 6). The *DOF* of the system is not fundamentally limited, but only restricted by the selected scanning mechanism and its range of lateral movement such that objects with depth range greater than 500  $\mu\text{m}$  can be measured with the system. The field of view is smaller than original requirements based on the size of the vision sensor. The optical resolution (depth) requirement is met, and lateral resolution is nearly met. The lateral resolution is slightly coarser than required since a 10X microscope is used rather than a 12X, due to availability.

The requirement detailing the number of images does not directly translate with the use of ROI and vision sensors outputting pixel-level events. Neither of these types of outputs are full frames. A total of two frames is listed in requirements since each pixel must be accessed twice, once for each scan.



Table 6. Comparison of final system design performance compared to requirements.

Item	Original Requirement	Updated Requirement	Design Performance
DOF ( $\mu\text{m}$ )	500	500	500+
FOV (mm x mm)	2.5 x 2.5	Less than 2.5 x 2.5	0.8 x 0.6
Optical Resolution (Depth) ( $\mu\text{m}$ )	3	3	3
Object Pixel Size ( $\mu\text{m}$ )	1.25	1.25	1.5
Number of Images	1 – 3	1 – 3	2
Grabbing Time per FOV (ms)	< 40	< 40	200+
Processing Time (ms)	Not specified	Not specified	Expected to be small compared to grabbing time
Estimated System Size (mm x mm x mm)	180 x 25 x 50	180 x 25 x 50	600 x 600 x 150
True Color Image?	Yes	Yes	Yes

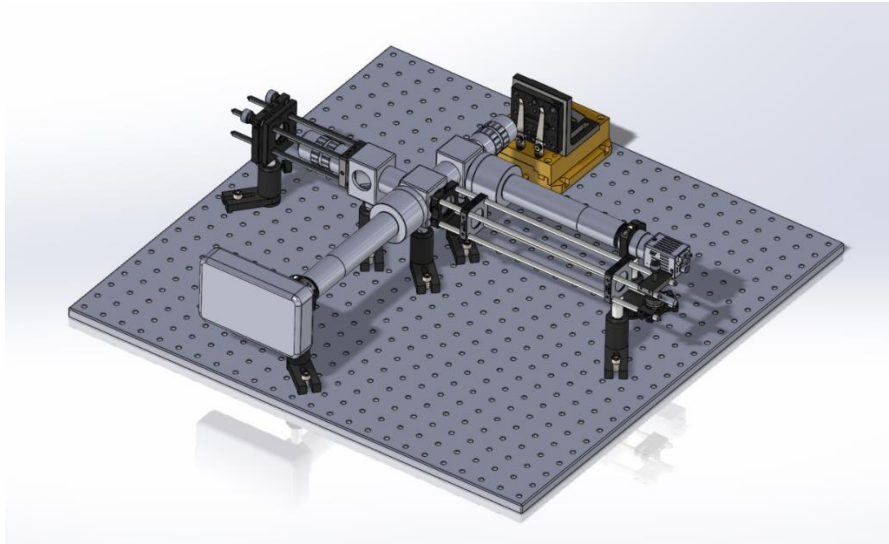
The speed of operation is not as fast as required; however, the limited speed comes from sensor capabilities rather than a limit in the system concept. This prototype – once built – will not operate as quickly as required, but it is the first step within the development process to meet all requirements. Speed is sacrificed in order to create a system that operates in a unique way and is well-suited to EDOF applications. Exact speeds cannot be quoted at this time since the system must work for a general object. Each object may require a unique number of ROIs of varying sizes and shapes, so different objects may be imaged at different speeds. Because of the slower “grabbing time” for the system, it is expected processing time will be shorter and not limit the system at this juncture.

The prototype system is larger than the requirements allow. This requirement is secondary to operation and performance, so allowing the size to increase somewhat to accommodate the use of two cameras is acceptable. In future iterations of the system, the size can be drastically reduced by folding the system, using integrated components, and by further refining the design. Therefore, the overall size is not considered for this prototype as the primary goals are based on system performance.

Finally, the system is able to create true color images and do so simply. There is no issue with processing interference fringes to gather true color information. The interferometer informs the imager, which is separate and shows no interference patterns, making true color imaging simpler rather than a main challenge of the system.



## 6.1 Future Work



*Figure 71. Layout of designed system with all components*

The prototype system (Figure 71) is to be built in the near future to confirm the system operates as expected. Components of the system have been purchased and building has begun. It is expected there will be some alignment challenges when building the system; however, this has been minimized by mounting optics using the prototype microscope scheme with C-Mount compatible components. These allow parts to be centered to one another while remaining in their designed positions. The challenges come from aligning the reference arm and test arm to these C-mounted parts, which is mitigated somewhat by cage mounting structures.

It is expected there will be some challenges with vibration from scanning the test object during operation. There is potential for lateral vibration which may blur the final image, obscuring fine details in the object.

The system requirements must be confirmed by testing the traditional optical performance (e.g., field of view and resolution) of the imaging system with a test chart. This is easy and straightforward, and it is not expected many issues will arise. Performance in depth must also be measured, which may involve 3D printing or machining of test parts to confirm depth resolution. Additionally, to confirm its capabilities in the field, the system will be tested with samples of devices that will be measured in practice.

The most significant challenge in future work is expected to be the creation of the algorithm to transform the input of the pixel-level depth map and produce optimized ROIs for the imaging sensor. Creating such an algorithm will be involved and rigorous. However, it is necessary to create the fastest operating system possible with the selected sensors. It also may be difficult to perfectly align the system to produce white light fringes, as the fringes are extremely sensitive to difference within the two interferometer arms.

## 6.2 Design Implications

The prototype lays a foundation for a new type of EDOF system. The system combines interferometry and imaging so the two operations do not significantly interfere with one another but, rather, work together to create an in-focus, true color image of a three-dimensional object, which has no significant artifacts and requires only limited processing. In addition, a coarse depth map is created in the process which may also inform the inspection process.

The system is limited in operation speed by currently available sensor technology. However, there have been great innovations in CMOS sensors in the past few years and the trend is expected to continue. CMOS sensors are likely to continue to implement processing at the pixel level and be able to do so on larger and larger sensors. It is likely that in the next few years, sensors necessary for this system will be able to operate quickly and the system will meet speed requirements. Additionally, on-chip processing may develop, allowing the system of Section 3.1 to be implemented. These developments in CMOS technology will likely remove a great number of challenges in this system and perhaps allow it even more flexibility in operation speed and in field of view.

The framework of the system allows for quick, detailed, and true color measurements of objects to be collected in a way other systems are not able to perform. Such a framework can be applied in-line inspection of many materials, such as semiconductor components, pharmaceuticals, and other small and detailed objects that require stringent quality control. The system can extend beyond in-line inspection as well, and it could be a valuable tool in many machine vision applications, medicinal applications (e.g., in-vivo inspection), and as a research tool in lab settings, especially for biological samples which must be captured quickly.

## References

- Abramowitz, Mortimer and Davidson, Michael W. "Kohler Illumination in Reflected Light." n.d. *Olympus*. 12 March 2019.
- Advanced Illumination. "MicroBrite Spot/Coaxial Light." 2017. Datasheet.
- Allied Vision. "Bonito Pro Technical Manual." 8 March 2019.
- Arecchi, Angelo V, Tahar Messadi and R. John Koshel. "Classical Illumination Designs." *Field Guide to Illumination*. Bellingham: SPIE Press, 2007.
- Ashok, Amit. 10 August 2018. Personal Interview.
- Blackman, Greg. "Stacked sensor with parallel processing built for fast, edge imaging." 19 February 2019. *Imaging and Machine Vision Europe*. 20 February 2019.
- Caber, Paul J. "Interferometric profiler for rough surfaces." *Applied Optics*, Vol. 32, No. 19 (1993).
- Castro, Albertina and Ojedga-Castaneda, Jorge. "Asymmetric phase masks for extended depth of field." *Applied Optics*, Vol. 43, No. 17 (2004).
- Chong, F. et al. "Optimization of spinning disk confocal microscopy: synchronization with the ultra-sensitive EMCCD." *Proc. SPIE 5324, Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XI*. 2004.
- Chouinard, Jon. "The Fundamentals of Camera and Image Sensor Technology." n.d. *AIA: Vision Online*. PowerPoint.
- Davies, Tom. "Coherence Modeling in FRED." Photon Engineering, August 2018. Email.
- de Groot, Peter. "Principles of interference microscopy for the measurement of surface topography." *Advances in Optics and Photonics* (2015).
- Demolder, Damien and Sanyal, Rishi. "Sensor breakthrough: Sony has developed a backlit CMOS sensor with global shutter." 16 February 2018. *Digital Photography Review*. 13 September 2018.
- "DLP Products: High speed visible - Products." 2019. *Texas Instruments*. Product Overview.
- Douglass, Michael R. *Digital Micromirror Device Reliability and Failure Mechanisms*. 25-27 October 2004. PowerPoint.
- Dumas, John P. et. al. "From modeling to hardware: an experimental evaluation of image plane and Fourier plane coded compressive optical imaging." *Optics Express*, Vol. 25, No. 23 (2017).
- Duocastella, M. et al. "Simultaneous imaging of multiple focal planes for three-dimensional microscopy using ultra-high-speed adaptive optics." *Journal of Biomedical Optics* (2012).
- Edmund Optics. "Infinity Corrected Objectives." n.d. *Edmund Optics*. 23 March 2019.
- Elouardi, A. et al. "Image Processing: towards a System on Chip." 1 December 2009. *IntechOpen*.
- Farber, V. et al. "Super-resolution compressive imaging with anamorphic optics." *Optics Express* Vol. 21, Issue 22 (2013).
- Fraunhofer IIS. "Our Services." *Industrial Image Processing*. n.d. Website. 24 March 2019.
- "FRED Application Note: Modeling Coherence." 2010. *Photon Engineering*. 12 March 2019.
- "Full Wave Rectifier." n.d. *Electronics Tutorials*. 12 March 2019.
- Gmuender, Tommy. *DLP Using Digital Micromirror Devices: A Primer*. SPIE Press, 2016.
- JAI. "Sequence Trigger Function." *User Manual: GO-2400M-PGE GO-2400C-PGE, 2.35M Digital Progressive Scan Monochrome and Color Camera*. October 2015.

- Jung, S. et al. "High-speed line scanning confocal microscope for biological imaging." *Proc. SPIE 6443, Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIV*. 2007.
- Kwon, Osuk, J.C. Wyant and C.R. and Hayslett. "Rough surface interferometry at 10.6  $\mu\text{m}$ ." *Applied Optics*, Vol. 19, No. 11 (1980).
- Lam, Edmund Y. "Computational photography with plenoptic camera and light field capture: tutorial." *Journal of the Optical Society of America A*, Vol. 32, Issue 11 (2015).
- Lee, Benjamin. "Introduction to +/- 12 Degree Orthogonal Digital Micromirror Devices (DMDs)." 2008-2018. *Texas Instruments*. 17 March 2019.
- Lichtsteiner, Patrick, Christoph Posch and Tobi Delbruck. "A 128 x 128 120 dB 15  $\mu\text{m}$  Latency Asynchronous Temporal Contrast Vision Sensor." *IEEE Journal of Solid-State Circuits*, Vol. 43, Issue 2. (2008).
- Mathew, Manoj V. "Kohler Illumination." 2015. *Lenstrotek*. 26 March 2019.
- Mueggler, E. et al. "The Event-Camera Dataset and Simulator: Event-based Data for Pose Estimation, Visual Odometry, and SLAM." *The International Journal of Robotics Research*, Volume 36, Issue 2 (2017).
- Ng, R. et al. "Light Field Photography with a Hand-held Plenoptic Camera." *Stanford Tech Report CTSR 2005-02* (2005).
- Nygren, Anders Johan. "Introduction." 8 May 1997. *Tomographic Reconstruction of SPECT*. 9 April 2019.
- Optronis. "CamRecord CR Series CR450x2(x3) CR600x2 CR1000x2(x3) CR3000x2 CR4000x2 CR5000x2 User Manual." n.d.
- Ouyang, Bing et. al. "Investigation of a compressive line sensing hyperspectral imaging sensor." *Proc. SPIE 10677, Unconventional Optical Imaging*. 2018.
- Pacheco, Shaun. "Array Confocal Microscopy." 2017.
- Pavliček, Pavel and Hýbl, Ondřej. "White-light interferometry on rough surfaces - measurement uncertainty caused by surface roughness." *Applied Optics*, Vol. 47, Issue 16 (2008): 2941-2949.
- Photon Engineering. "Coherent Sources Overview." *FRED Help*. n.d.
- . "Surface Roughness - Sampled Random Surface Normal." *FRED Help*. n.d.
- "Power Diodes and Rectifiers." n.d. *Electronics Tutorials*. 12 March 2019.
- Prophesee. "Prophesee Evaluation Kit." *The First Event-Based Vision System*. n.d.
- . "Prophesee Metavision for Machines." *White Paper: Event-based sensing enables a new generation of machine vision solutions*. n.d.
- Roveda, Janet. 7 August 2018. Personal Interview.
- Sasian, Jose. "OPTI 696A Lecture." 4 December 2018.
- Schmit, Joanna. "3D white light interference microscope with specialized illumination for better sample imaging and observation." *Advanced Mechatronics Solutions, Advances in Intelligent Systems and Computing* (2016).
- . "An introduction to non-contact surface metrology." 2013. *Bruker*. PowerPoint. 16 March 2019.
- Schmit, Joanna, Matt Novak and Son and Bui. "3D interferometric microscope: color visualization of engineered surfaces for industrial applications." *Applied Advanced Optical Metrology Solutions*. Ed. James D. Trolinger Erik Novak. SPIE, 2015.
- Schrey, O. et al. "A 1 K/spl times/1 K high dynamic range CMOS image sensor with on-chip programmable region-of-interest readout." *IEEE Journal of Solid-State Circuits* (2002).

Stemmer Imaging. *Machine vision basics: Angles of illumination*. n.d. Video Tutorial.

—. "Region of interest (ROI)." n.d. *Stemmer Imaging*. Knowledge Base Article. 24 March 2019.

"The Differential Amplifier." n.d. *Electronics Tutorials*. 12 March 2019.

Wyant, James C. "5.0 Direct Phase Measurement Interferometry." 5 November 2012. *OPTI 513: Optical Testing*. PowerPoint. 13 March 2019.

Wyant, James C. and Katherine Creath. "Advances in Interferometric Optical Profiling." *International Journal of Machine Tools and Manufacture* (1992).

Yang, Shuming and Zhang, Guofeng. "A review of interferometry for geometric measurement." *Measurement Science and Technology* (Volume 29, Number 10, 2018).

Zhang, Yu et. al. "Light Field Modeling for Coded Aperture Systems." *Imaging and Applied Optics* (2016).